

**This Page Is Inserted by IFW Operations
and is not a part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SLANTED IMAGES**
- **COLORED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: 11234668 A

(43) Date of publication of application: 27 . 08 . 99

(51) Int. Cl.

H04N 7/24

(21) Application number: 10031614

(22) Date of filing: 13 . 02 . 98

(71) Applicant: SONY CORP

(72) Inventor: MATSUNAGA OSAMU
FUJISHIRO SHIGEO

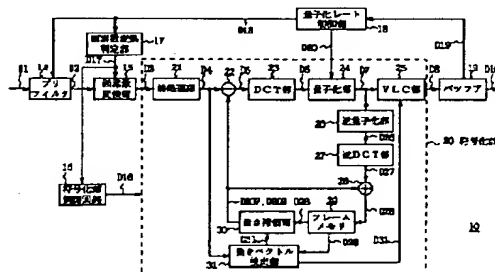
(54) VIDEO CODER

(57) Abstract:

PROBLEM TO BE SOLVED: To provide the video coder that conducts band reduction processing and pixel number conversion processing in response to a compressing difficulty of an image.

SOLUTION: The coder is provided with a pre-filter 14 that reduces a signal band of a video signal and a pixel number conversion section 15 that converts a pixel number of the video signal. The pixel number conversion section 15 applies pixel number conversion adaptively to a video signal so as to have a pixel number in response to a signal band reduced by the pre-filter 14. So that a generated code quantity is reduced in the coding processing, the quantization accuracy is improved in the quantization processing and image quality deterioration is suppressed in the coding processing.

COPYRIGHT: (C)1999,JPO



BEST AVAILABLE COPY

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-234668

(43) 公開日 平成11年(1999) 8月27日

(51) Int.Cl.⁶

識別記号

F I

H 0 4 N 7/24

H 0 4 N 7/13

Z

審査請求 未請求 請求項の数 6 O L (全 8 頁)

(21) 出願番号 特願平10-31614

(22) 出願日 平成10年(1998) 2月13日

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川 6 丁目 7 番 35 号

(72) 発明者 松永 修

東京都品川区北品川 6 丁目 7 番 35 号 ソニー
株式会社内

(72) 発明者 藤代 茂夫

東京都品川区北品川 6 丁目 7 番 35 号 ソニー
株式会社内

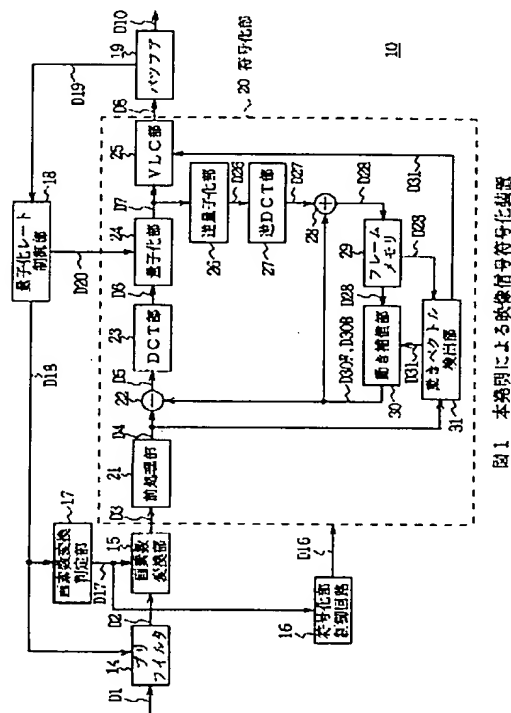
(74) 代理人 弁理士 田辺 恵基

(54) 【発明の名称】 映像符号化装置

(57) 【要約】

【課題】画像の圧縮難易度に応じた帯域削減処理及び画素数変換処理を行う映像符号化装置を提案する。

【解決手段】映像信号の信号帯域を削減するプリフィルタ 1 4 と、映像信号の画素数を変換する画素数変換部 1 5 とを設け、当該画素数変換部 1 5 において当該プリフィルタにより削減された信号帯域に応じた画素数に適応的に画素数変換を行うようにしたことにより、符号化処理による発生符号量を減少させ、量子化処理における量子化精度を向上し、符号化処理による画質劣化を抑え得る。



【特許請求の範囲】

【請求項 1】 入力映像信号を圧縮符号化して送出する映像符号化装置において、

上記入力映像信号の信号帯域を削減する信号帯域削減手段と、

上記削減された信号帯域に基づいて上記入力映像信号の画素数を変換する画素数変換手段と、

上記画素数変換手段により画素数を変換された上記入力映像信号を圧縮符号化する符号化手段とを具えることを特徴とする映像符号化装置。

【請求項 2】 上記信号帯域削減手段は、

上記入力映像信号の符号化難易度に基づいて上記入力映像信号の信号帯域を削減することを特徴とする請求項 1 に記載の映像符号化装置。

【請求項 3】 上記信号帯域削減手段は、複数の通過帯域幅を設定する水平周波数低域通過型フィルタであることを特徴とする請求項 1 に記載の映像符号化装置。

【請求項 4】 上記画素数変換手段は、上記映像信号の水平方向画素数について画素数変換を行うことを特徴とする請求項 1 に記載の映像符号化装置。

【請求項 5】 上記画素数変換手段は、

上記信号帯域を削減された上記入力映像信号を表現するに必要な画素数を判定し、当該判定結果を上記変換画素数として設定する画素数変換判定手段を具えることを特徴とする請求項 1 に記載の映像符号化装置。

【請求項 6】 上記画素数変換手段は、

上記入力映像信号のフレーム同期信号に基づくタイミングで上記画素数を変換することを特徴とする請求項 1 に記載の映像符号化装置。

【発明の詳細な説明】

【0001】

【目次】 以下の順序で本発明を説明する。

【0002】 発明の属する技術分野

従来の技術（図 4）

発明が解決しようとする課題

課題を解決するための手段

発明の実施の形態（図 1～図 3）

発明の効果

【0003】

【発明の属する技術分野】 本発明は映像符号化装置に関し、例えばデジタル放送に用いられる映像信号符号化装置に適用して好適なものである。

【0004】

【従来の技術】 近年、映像及び音声の情報量を減らす方法として種々の圧縮符号化方法が提案されており、その代表的なものに M P E G 2 (Moving Picture Experts Group Phase 2) と呼ばれる方式がある。かかる M P E G 2 方式を用いて、映像及び音声の放送データを圧縮符号化し、地上波または衛星波を用いて放送するデジタル放送システムが開始されている。

【0005】 図 4 は全体として映像信号符号化装置 5 0 を示し、ビデオテープレコーダ等の映像信号供給装置

（図示せず）から映像信号 D 1 がプリフィルタ 5 1 に入力されている。映像信号 D 1 はプリフィルタ 5 1 において、量子化レート制御部 5 5 から供給される周波数特性制御信号 D 5 5 に応じた帯域制限処理を受けることにより映像信号の高域成分が削減され、帯域制限映像信号 D 5 1 として画素数変換部 5 2 に出力される。ここで、映像信号の高域成分は当該映像の細かい部分を表現するものであり、当該高域成分を削減することにより映像の細かい部分を省略し、映像全体に対する影響を少なく留めつつ映像信号の帯域幅を減少させる。

【0006】 画素数変換部 5 2 は、入力された帯域制限映像信号 D 5 1 に対し画素数変換処理を行う。すなわち符号化装置 5 0 に入力された映像信号 D 1 の水平方向の画素数を水平画素数 M とすると、画素数変換部 5 2 は、当該映像信号 S 1 を帯域制限してなる帯域制限映像信号 D 5 1 の水平方向画素数を、 $N < M$ である削減水平画素数 N に削減し、画素数変換映像信号 D 5 2 として符号化部 5 3 に送出する。ここで削減水平画素数 N は、映像信号 D 1 の番組内容に応じて、高画質を要求される番組では大きく、高画質を要求されない番組では小さく設定される。

【0007】 画素数変換映像信号 D 5 2 は、符号化部 5 3 において動き補償処理、D C T (Discrete Cosine Transform、離散コサイン変換) 処理、量子化処理、及び V L C (Variable Length Coding、可変長符号化) 処理を受け圧縮符号化され、可変長符号化データ D 5 3 としてバッファ 5 4 に送出される。このとき符号化部 5 3

は、量子化レート制御部 5 5 から供給される量子化制御信号 D 5 6 に応じて量子化処理における量子化レートを調整する。また、符号化部 5 3 には符号化部制御回路 5 7 から符号化部制御情報 D 5 7 が供給されており、当該符号化部 5 3 は当該符号化部制御情報 D 5 7 に応じて符号化タイミング及び動き補償処理における動きベクトル探索範囲を設定する。

【0008】 ここで、映像信号 D 1 の映像の動き量が多い場合や当該映像が複雑な場合、当該映像信号 D 1 を圧縮符号化してなる可変長符号化データ D 5 3 の発生符号量は増大する。このような圧縮符号化による符号発生量の多い映像信号を、圧縮難易度が高いと表現する。かかる圧縮難易度は、映像信号の番組内容やシーンに応じて変動する。このため、バッファ 5 4 における可変長符号化データ D 5 3 の占有量を基に、符号化部 5 3 における量子化処理の量子化値及びプリフィルタ 5 1 における帯域制限を制御することにより、一定期間例えば G O P (Group Of Pictures) 当たりの発生符号量を一定に制御する。

【0009】 すなわち、量子化レート制御部 5 5 はバッファ 5 4 における可変長符号化データ D 5 3 の蓄積状態

を常に監視しており、かかる蓄積状態を占有率情報D 5 4として得る。そして、量子化レート制御部5 5は占有率情報D 5 4を基に量子化制御信号D 5 6及び周波数特性制御信号D 5 5を生成し、それぞれを符号化部5 3及びプリフィルタ5 1に供給することにより、可変長符号化データD 5 3の一定期間当たりの発生符号量を一定に制御するようになされている。

【0010】

【発明が解決しようとする課題】ところで、かかる映像信号符号化装置5 0においては、プリフィルタ5 1による帯域削減幅は映像信号D 1の圧縮難易度によつて変動するのに対し、画素数変換部5 2による削減水平画素数Nは映像信号D 1の番組内容に応じて一定である。このため、画素数削減映像信号D 5 2の水平方向画素数Nが、帯域制限映像信号D 5 1の映像を表現するに必要な水平方向画素数を上回る場合が考えられる。この場合、帯域制限映像信号D 5 1の映像を表現するに必要な画素数以上の画素を符号化することになり、量子化処理における量子化値を必要以上に増やすことになり画質劣化を引き起こすという問題があつた。

【0011】本発明は以上の点を考慮してなされたもので、画像の圧縮難易度に応じた帯域削減処理及び画素数変換処理を行う映像符号化装置を提案しようとするものである。

【0012】

【課題を解決するための手段】かかる課題を解決するため本発明においては、映像信号の信号帯域を削減する信号帯域削減手段と、映像信号の画素数を変換する画素数変換手段とを設け、当該画素数変換手段において当該信号帯域削減手段により削減された信号帯域に応じた画素数に適応的に画素数変換を行うようにする。

【0013】

【発明の実施の形態】以下図面について、本発明の一実施の形態を詳述する。

【0014】図1は全体として映像信号符号化装置1 0を示し、ビデオテープレコーダ等の映像信号供給装置(図示せず)からHDTV (High Definition Television、高精細度テレビジョン)規格の映像信号D 1がプリフィルタ1 4に入力されている。

【0015】プリフィルタ1 4は水平周波数低域通過型フィルタであり、映像信号D 1に対し、量子化レート制御部1 8から供給される周波数特性制御信号D 1 8に応じて当該映像信号D 1の高域成分を削減する帯域制限処理を行い、映像の細かい部分を省略し映像全体に対する影響を少なく留めつつ当該映像信号の帯域幅を減少し、帯域制限映像信号D 2として画素数変換部1 5に出力する。

【0016】画素数変換部1 5は、入力された帯域制限映像信号D 2に対し、画素数変換判定部1 7から供給される画素数制御信号D 1 7に応じて当該帯域制限映像信

号D 2の水平方向画素数を削減し、画素数変換映像信号D 3として符号化部2 0に送出する。例えば、符号化装置1 0に入力された映像信号D 1の水平方向の画素数を1 9 2 0画素とすると、画素数変換部1 5は、当該映像信号D 1を帯域制限してなる帯域制限映像信号D 2の水平方向画素数を、 $N < 1 9 2 0$ である削減水平画素数Nに削減し、これを画素数変換映像信号D 3として符号化部2 0に送出する。

【0017】この場合、画素数変換判定部1 7は、量子化レート制御部1 8から送出される周波数特性制御信号D 1 8に応じて画素数制御信号D 1 7を生成する。まず画素数変換判定部1 7は、周波数特性制御信号D 1 8が示す高周波成分削減処理後の帯域幅を基に、帯域制限映像信号D 2の映像を表現するに必要な水平方向画素数である必要水平画素数N'を求める。周波数特性制御信号D 1 8の示す帯域幅をF [Hz]とし、帯域制限映像信号D 2の水平走査時間をT h [秒]とすると、 $N' = 2 F \cdot T h$ [画素]となる。

【0018】次に画素数変換判定部1 7は、予め設定してある複数の設定水平画素値N sから、 $N' < N s$ の条件を満たす最も小さい値の設定水平画素値N sを選択し、これを削減水平画素数Nとする。この実施の形態においては、設定水平画素値N sは1 4 4 0画素、1 2 8 0画素、1 1 5 2画素、9 6 0画素の4種類が設定されている。そして画素数変換判定部1 7は、各設定水平画素値に対応して設定された識別番号を画素数制御信号D 1 7として画素数変換部1 5及び符号化部制御回路1 6に送出する。ここで設定水平画素値の種類をj種類とすると、画素数制御信号D 1 7はkビット($j \leq 2^k$)の信号で表せられる。

【0019】かくして帯域制限映像信号D 2は、画素数変換部1 5において当該帯域制限映像信号D 2を表現するに十分な削減水平画素数Nに画素数変換され、画素数変換映像信号D 3として符号化部2 0に送出される。

【0020】また、符号化部制御回路1 6は、画素数変換値D 1 7を基に符号化部制御信号D 1 6を符号化部2 0に供給する。符号化部2 0は、符号化部制御信号D 1 6に応じて動作することにより、画素数変換映像信号D 3の水平画素数に応じた符号化処理を行う。すなわち、図2 (A)に示す水平画素数Mの帯域制限映像信号D 2は、画素数変換部1 5において画素数変換され、図2 (B)に示す水平画素数Nの画素数変換映像信号D 3として符号化部2 0に送出される。ここで、画素数変換映像信号D 3は連続して送出されるわけではなく、図2 (B)に示すようにN画素が連続して送出された後、(M-N)画素に相当する期間は送出が停止する。符号化部2 0は、符号化部制御信号D 1 6に応じてN画素に相当する期間は符号化動作を行い、(M-N)画素に相当する期間は符号化動作を停止することにより、動作クロックを一定に保ちつつ、画素数変換映像信号D 3の水

平面素数に応じた符号化処理を行う。また符号化部 2 0 は、符号化部制御信号 D 1 6 に応じて動き補償処理における動きベクトル探索範囲を設定する。

【0021】符号化部 2 0 (図 1)において、画素数変換映像信号 D 3 は前処理部 2 1 に入力される。前処理部 2 1 は、順次入力される画素数変換映像信号 D 3 の各フレーム画像について I ピクチャ、P ピクチャまたは B ピクチャの 3 つの画像タイプのうちのどの画像タイプとして処理するかを指定した後、当該フレーム画像の画像タイプに応じて当該フレーム画像を符号化する順番に並べ替え、さらに当該フレーム画像を 1 6 画素×1 6 ラインの輝度信号及び当該輝度信号に対応する色差信号によって構成されるマクロブロックに分割し、これをマクロブロックデータ D 4 として、演算回路 2 2 及び動きベクトル検出部 3 1 に供給する。

【0022】動きベクトル検出部 3 1 は、マクロブロックデータ D 4 の各マクロブロックの動きベクトルを、当該マクロブロックデータ D 4 及びフレームメモリ 2 9 に記憶されている参照画像データ D 2 8 を基に算出し、動きベクトルデータ D 3 1 として動き補償部 3 0 及び V L C (Variable Length Coding、可変長符号化)部 2 5 に送出する。

【0023】演算回路 2 2 は、前処理部 2 1 から供給されたマクロブロックデータ D 4 について、当該マクロブロックデータ D 4 の各マクロブロックの画像タイプに基づいて、イントラモード、順方向予測モード、逆方向予測モードまたは双方向予測モードのいずれかの予測モードの動き補償を行う。ここでイントラモードとは、符号化対象となるフレーム画像をそのまま伝送データとする方法であり、順方向予測モードとは、符号化対象となるフレーム画像と過去参照画像との予測残差を伝送データとする方法である。また逆方向予測モードとは、符号化対象となるフレーム画像と未来参照画像との予測残差を伝送データとする方法であり、双方向予測モードとは、符号化対象となるフレーム画像と、過去参照画像及び未来参照画像の 2 つの予測画像の平均値との予測残差を伝送データとする方法である。

【0024】まず、マクロブロックデータ D 4 が I ピクチャである場合について説明する。この場合、マクロブロックデータ D 4 はイントラモードで処理される。すなわち、演算回路 2 2 はマクロブロックデータ D 4 のマクロブロックを、そのまま演算データ D 5 として D C T

(Discrete Cosine Transform、離散コサイン変換)部 2 3 に送出する。D C T 部 2 3 は演算データ D 5 に対し D C T 変換処理を行い D C T 係数化し、D C T 係数データ D 6 として量子化部 2 4 に送出する。量子化部 2 4 は D C T 係数データ D 6 に対し量子化処理を行い、量子化 D C T 係数データ D 7 として V L C 部 2 5 及び逆量子化部 2 6 に送出する。このとき量子化部 2 4 は、量子化レート制御部 1 8 より供給される量子化制御信号 D 2 0 に

応じて量子化値を調整することにより、発生する符号量が制御される。

【0025】逆量子化部 2 6 に送出された量子化 D C T 係数データ D 7 は逆量子化処理を受け、D C T 係数データ D 2 6 として逆 D C T 部 2 7 に送出される。そして D C T 係数データ D 2 6 は、逆 D C T 部 2 7 において逆 D C T 処理を受け、演算データ D 2 7 として演算回路 2 8 に送出され、参照画像データ D 2 8 としてフレームメモリ 2 9 に記憶される。

【0026】次に、マクロブロックデータ D 4 が P ピクチャである場合について説明する。この場合、演算回路 2 2 はマクロブロックデータ D 4 について、イントラモードまたは順方向予測モードのいずれかの予測モードによる動き補償処理を行う。

【0027】予測モードがイントラモードの場合、上述の I ピクチャの場合と同様に、演算回路 2 2 はマクロブロックデータ D 2 のマクロブロックをそのまま演算データ D 5 として D C T 部 2 3 に送出する。

【0028】これに対して、予測モードが順方向予測モードの場合、演算回路 2 2 はマクロブロックデータ D 4 について、動き補償部 3 0 より供給される順方向予測画像データ D 3 0 F を用いて減算処理する。

【0029】順方向予測画像データ D 3 0 F は、フレームメモリ 2 9 に記憶されている参照画像データ D 2 8 を、動きベクトルデータ D 3 1 に応じて動き補償することにより算出される。すなわち動き補償部 3 0 は順方向予測モードにおいて、フレームメモリ 2 9 の読出アドレスを動きベクトルデータ D 3 1 に応じてずらして参照画像データ D 2 8 を読み出し、これを順方向予測画像データ D 3 0 F として演算回路 2 2 及び演算回路 2 8 に供給する。演算回路 2 2 はマクロブロックデータ D 4 から順方向予測画像データ D 3 0 F を減算して予測残差としての差分データを得、演算データ D 5 として D C T 部 2 3 に送出する。

【0030】また、演算回路 2 8 には動き補償部 3 0 から順方向予測画像データ D 3 0 F が供給されており、演算回路 2 8 は演算データ D 2 7 に当該順方向予測画像データ D 3 0 F を加算することにより参照画像データ D 2 8 (P ピクチャ)を局部再生し、フレームメモリ 2 9 に記憶する。

【0031】次に、前処理部 2 1 から B ピクチャのマクロブロックデータ D 4 が演算回路 2 2 に供給された場合について説明する。この場合、演算回路 2 2 はマクロブロックデータ D 4 について、イントラモード、順方向予測モード、逆方向予測モードまたは双方向予測モードのいずれかの動き補償処理を行う。

【0032】予測モードがイントラモードまたは順方向予測モードの場合、マクロブロックデータ D 4 は上述の P ピクチャの場合と同様の処理を受ける。但し、B ピクチャは他の予測参照画像として用いられないので、参照

画像データ D 2 8 はフレームメモリ 2 9 には記憶されない。

【0 0 3 3】これに対して、予測モードが逆方向予測モードの場合、演算回路 2 2 はマクロブロックデータ D 4 について、動き補償部 3 0 より供給される逆方向予測画像データ D 3 0 B を用いて減算処理する。

【0 0 3 4】逆方向予測画像データ D 3 0 B は、フレームメモリ 2 9 に記憶されている参照画像データ D 2 8 を、動きベクトルデータ D 3 1 に応じて動き補償することにより算出される。すなわち動き補償部 3 0 は逆方向予測モードにおいて、フレームメモリ 2 9 の読出アドレスを動きベクトルデータ D 3 1 に応じてずらして参照画像データ D 2 8 を読み出し、これを逆方向予測画像データ D 3 0 B として演算回路 2 2 及び演算回路 2 8 に供給する。演算回路 2 2 はマクロブロックデータ D 4 から逆方向予測画像データ D 3 0 B を減算して予測残差としての差分データを得、演算データ D 5 として D C T 部 2 3 に送出する。

【0 0 3 5】また、演算回路 2 8 には動き補償部 3 0 より逆方向予測画像データ D 3 0 B が供給されており、演算回路 2 8 は演算データ D 2 7 に当該逆方向予測画像データ D 3 0 B を加算することにより参照画像データ D 2 8 (B ピクチャ) を局部再生するが、B ピクチャは他の予測参照画像として用いられないので、参照画像データ D 2 8 はフレームメモリ 2 9 には記憶されない。

【0 0 3 6】予測モードが双方向モードの場合、演算回路 2 2 はマクロブロックデータ D 4 から、動き補償部 3 0 より供給される順方向予測画像データ D 3 0 F 及び逆方向予測画像データ D 3 0 B の平均値を減算し予測残差としての差分データを得、演算データ D 5 として D C T 部 2 3 に送出する。

【0 0 3 7】また、演算回路 2 8 には動き補償部 3 0 より順方向予測画像データ D 3 0 F 及び逆方向予測画像データ D 3 0 B が供給されており、演算回路 2 8 は演算データ D 2 7 に当該順方向予測画像データ D 3 0 F 及び逆方向予測画像データ D 3 0 B の平均値を加算することにより参照画像データ D 2 8 (B ピクチャ) を局部再生するが、B ピクチャは他の予測参照画像として用いられないので、参照画像データ D 2 8 はフレームメモリ 2 9 には記憶されない。

【0 0 3 8】かくして、符号化部 2 0 に入力された画素数変換映像信号 D 3 は、動き補償処理、D C T 処理及び量子化処理を受け、量子化 D C T 係数データ D 7 として V L C 部 2 5 に供給される。

【0 0 3 9】V L C 部 2 5 は、量子化 D C T 係数データ D 7 及び動きベクトルデータ D 3 1 に対し、所定の変換テーブルに基づく可変長符号化処理を行い、その結果得られるデータを可変長符号化データ D 8 としてバツファ 1 9 に送出する。バツファ 1 9 において可変長符号化データ D 8 は一時蓄積された後、順次可変長符号化データ

D 1 0 として読み出される。

【0 0 4 0】量子化レート制御部 1 8 は、バツファ 1 9 における可変長符号化データ D 8 の蓄積量を常に監視しており、かかる蓄積量を占有率情報 D 1 9 として得る。そして、量子化レート制御部 1 8 は占有率情報 D 1 9 を基に周波数特性制御信号 D 1 8 及び量子化制御信号 D 2 0 を生成し、それぞれをプリフィルタ 1 4 及び量子化部 2 4 に送出することにより、帯域削減処理における帯域幅及び量子化処理における量子化値を調整する。

【0 0 4 1】ここで、画素数変換映像信号 D 3 は映像信号 D 1 に比べて画素数が削減されているため、当該画素数変換映像信号 D 3 の画像を 1 6 画素×1 6 ラインに分割してなるマクロブロックデータ D 4 のマクロブロック数も減少する。動き補償処理における動きベクトルの合計量はマクロブロック数にほぼ比例するため、マクロブロックデータ D 4 のマクロブロック数が減少すると、これに伴い動きベクトルデータ D 3 1 の発生符号量も減少する。このため、D C T 係数データ D 7 及び動きベクトルデータ D 3 1 を可変長符号化処理して得られる可変長符号化データ D 8 の発生符号量は減少し、これに伴いバツファ 1 9 における可変長符号化データ D 8 の蓄積量も減少する。

【0 0 4 2】このため可変長符号化データ D 1 0 の発生符号量を一定とした場合、量子化レート制御部 1 8 は、バツファ 1 9 における可変長符号化データ D 8 の蓄積量の減少に応じて量子化部 1 8 における量子化値を細かく制御することにより、マクロブロック数の減少による動きベクトルデータ D 3 1 の発生符号量の減少に応じて D C T 係数データ D 6 をより細かい量子化値で量子化し得るため、量子化精度を向上させることができる。

【0 0 4 3】以上の構成において、映像信号符号化装置に入力された映像信号 D 1 は、プリフィルタ 1 4 において帯域削減処理を受け、帯域制限映像信号 D 2 として画素数変換部 1 5 に送出される。このとき量子化レート制御部 1 8 は、当該映像信号 D 1 の圧縮難易度に応じてプリフィルタ 1 4 における帯域削減量を制御する。

【0 0 4 4】帯域制限映像信号 D 2 は、画素数変換部 1 5 において、当該帯域制限映像信号 D 2 を表現するに必要な最小限な画素数に画素数削減を受け、画素数削減映像信号 D 3 として符号化部 2 0 に送出される。このとき、画素数の削減に伴いマクロブロックの数も減少する。

【0 0 4 5】画素数削減映像信号 D 3 は、符号化部 2 0 において動き補償処理、D C T 変換処理、量子化処理及び可変長符号化処理を受け、可変長符号化データ D 8 としてバツファ 1 9 に出力される。このとき量子化レート制御部 1 8 は、当該映像信号 D 1 の圧縮難易度に応じて符号化部 2 0 における量子化処理の量子化値を制御する。

【0 0 4 6】ここで、画素数削減映像信号 D 3 は帯域制限映像信号 D 2 に比べてマクロブロック数が減少してい

る。動き補償処理による動きベクトルの発生符号量はマクロブロック数にほぼ比例するため、当該画素数削減映像信号D 3を符号化してなる可変長符号化データD 8の動きベクトルの発生符号量も減少する。この動きベクトルの発生符号量が減少する分、DCT係数をより細かい量子化値で量子化できるため、画質劣化が回避できる。

【0047】以上の構成によれば、映像信号の信号帯域を削減するプリフィルタと、映像信号の画素数を変換する画素数変換部とを設け、当該画素数変換部において当該プリフィルタにより削減された信号帯域に応じた画素数に適応的に画素数変換を行い、必要最小限の画素を符号化することにより、量子化処理における量子化値が不必要に大きくなる事を防ぎ、符号化処理による画質劣化を防止し得る。

【0048】なお上述の実施の形態においては、画素数変換判定部17は量子化レート制御部18から送出される周波数特性制御信号D18に応じて画素数制御信号D17を送出するようにしたが、本発明はこれに限らず、映像信号D1のフレームを検出し、フレーム周期に同期して画素数制御信号D17を送出するようにしてもよい。すなわち、図1との同一部分に同一符号を付して示す図3において、11は全体として符号化装置を示し、図示しない映像信号供給装置から映像信号D1がプリフィルタ14及び同期信号生成部35に供給されている。同期信号生成部35は、映像信号D1のフレーム周期を検出し、当該フレーム周期に同期した同期信号D35を画素数変換判定部36に送出する。画素数変換判定部36は、画素数変換部15に対し、同期信号D35に同期して画素数制御信号D17を送出する。画素数変換部15は画素数制御信号D17に応じて画素数変換処理を行うことにより、帯域制限映像信号D2に対しフレーム単位での画素数変換処理を行い得る。

【0049】なお上述の実施の形態においては、画素数変換処理における変換水平方向画素数を1440画素、1280画素、1152画素及び960画素としたが、

本発明はこれに限らず、これ以外の水平方向画素数を用いるようにしても良い。

【0050】また上述の実施の形態においては、映像信号D1をHDTV規格の映像信号としたが、本発明はこれに限らず、HDTV以外の規格の映像信号を入力するようにしても良い。

【0051】また上述の実施の形態においては、画像の圧縮難易度に応じて水平方向画素数の削減処理を行うようにしたが、本発明はこれに限らず、垂直方向すなわち走査線本数を削減するようにしたり、垂直方向及び水平方向すなわち走査線本数及び水平方向画素数を同時に削減するようにしても良い。

【0052】

【発明の効果】上述のように本発明によれば、映像信号の信号帯域を削減する信号帯域削減手段と、映像信号の画素数を変換する画素数変換手段とを設け、当該画素数変換手段において当該信号帯域削減手段により削減された信号帯域に応じた画素数に適応的に画素数変換を行うようにしたことにより、符号化処理による発生符号量を減少させ、量子化処理における量子化精度を向上し、符号化処理による画質劣化を抑え得る。

【図面の簡単な説明】

【図1】本発明による映像信号符号化装置を示すブロック図である。

【図2】画素数削減を示す略線図である。

【図3】他の実施の形態による映像信号符号化装置を示すブロック図である。

【図4】映像信号符号化装置を示すブロック図である。

【符号の説明】

10、11、50……映像信号符号化装置、14、51……プリフィルタ、15、52……画素数変換部、16、57……符号化部制御回路、17、36……画素数変換判定部、18、55……量子化レート制御部、19、54……バッファ、20、53……符号化部、35……同期信号生成部。

【図2】

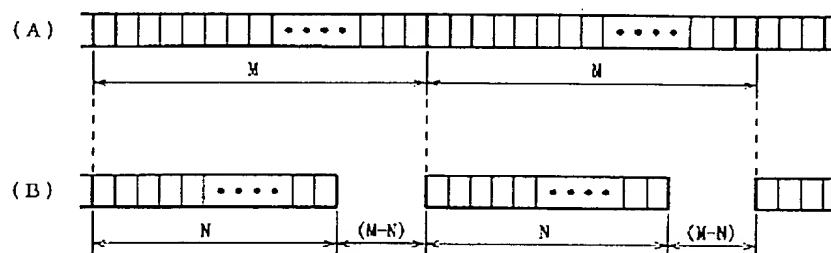


図2 画素数変換

【図 1】

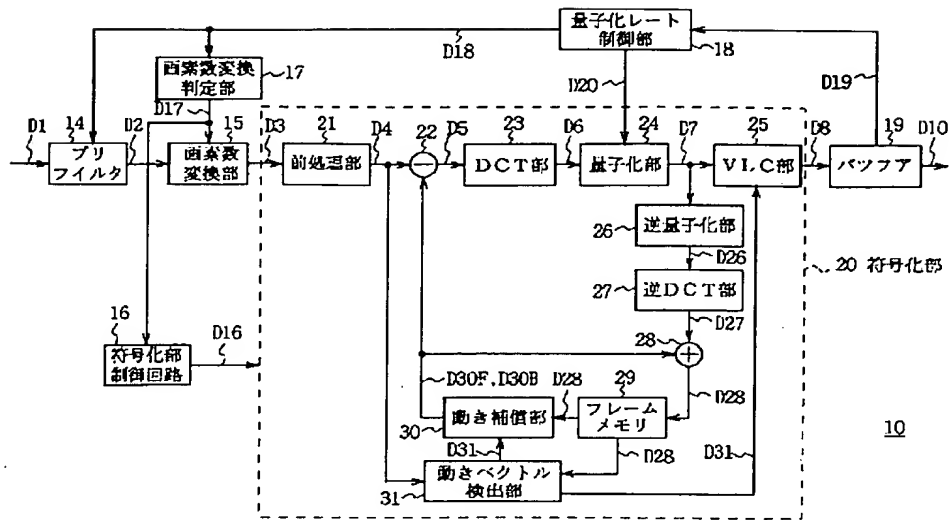


図 1 本発明による映像信号符号化装置

【図 3】

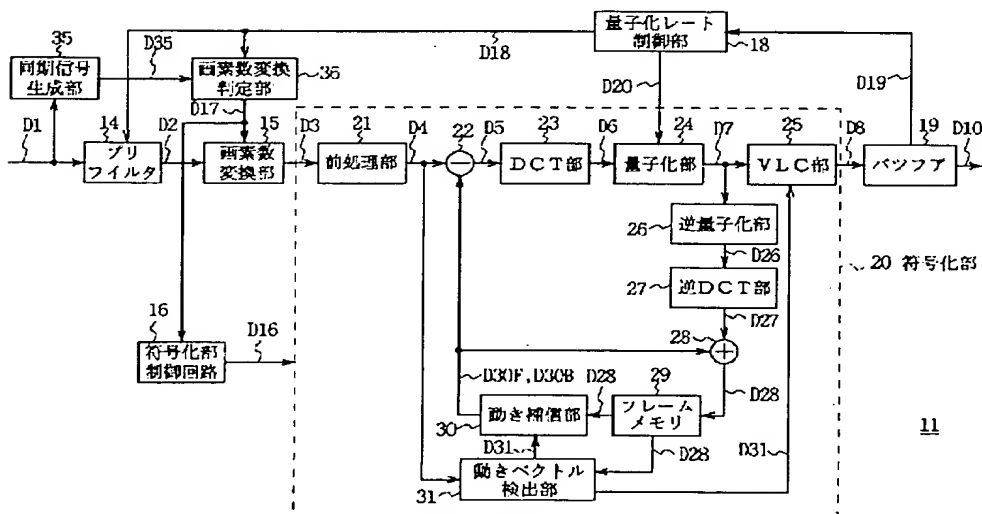


図 3 他の実施の形態による映像信号符号化装置

【図 4】

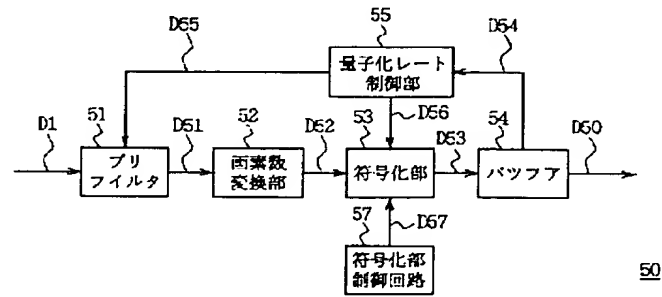


図 4 映像信号符号化装置

まず、4:2:0では、色差信号は水平・垂直方向ともにサブサンプルされて輝度信号の1/2のサイズになっており、図4(a)に示すような空間的なサンプリングを行っている。

4:2:2では図4(b)のようにサンプリング周波数の比が輝度に対して色差が1/2になっており、水平方向のみが半分のサイズになっている。

4:4:4では図4(c)のように輝度信号と色差信号のサンプリング周波数が全く同じ場合である。

1.4 アスペクト比

MPEG ではさまざまな応用分野を想定しているため、画像のアスペクト比は標準テレビの4:3、EDTV や HDTV の16:9、コンピュータ用に画素アスペクト比1:1を用いることができる。

2. シーケンス〜ブロック構造

MPEG のビデオの符号化ビットストリームの構成を図5に、また、MPEG-2でのビットストリームのブロックダイアグラムと主な内容をそれぞれ図6、表1に示す。

MPEG-1、MPEG-2 いずれも、符号化されたデータ(ビットストリーム)は階層構造になっており、上位層から順にシーケンス(Sequence)層、GOP(Group Of Picture)層、ピクチャ(Picture)層、スライス(Slice)層、マクロブロック(Macro Block)層、およびブロック(Block)層から構成される。シーケンス層からスライス層までの各階層では、ユニークなスタートコードが与えられており、これにより各層が区別されるとともに、エラー回復ポイントにもなっている。以下に MPEG-2 の場合の符号化ビットストリームの

構成を説明する。

2.1 シーケンス層

シーケンス層は一番上位に位置する階層で、シーケンスヘッダで始まり、図6のシーケンス拡張を示す extension start code の有無で MPEG-1 と MPEG-2 のビットストリームの分類が行われる。一般にシーケンスは1つのビデオプログラム全体を示し、sequence end code で終了する。また、シーケンス層は図5のように基本的に1個以上の GOP から構成される。

シーケンスヘッダには表1(a)のように、符号化画像サイズ、アスペクト比、フレームレート、ビットレート、VBV バッファサイズ、量子化マトリクスなど

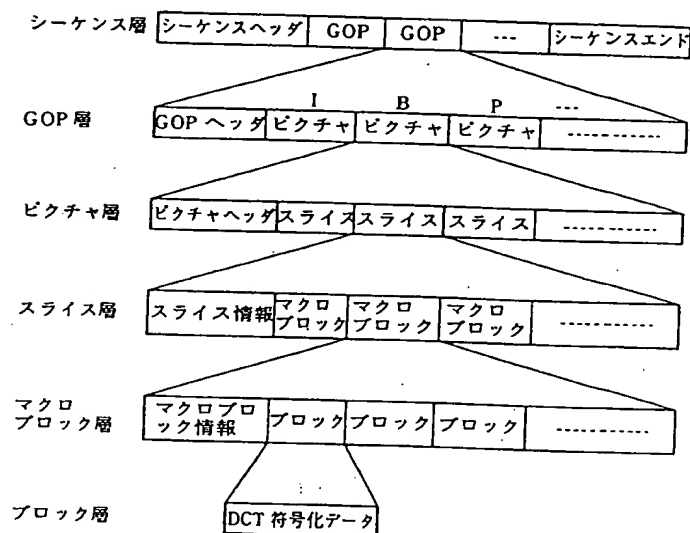


図5 MPEG ビットストリームの構成

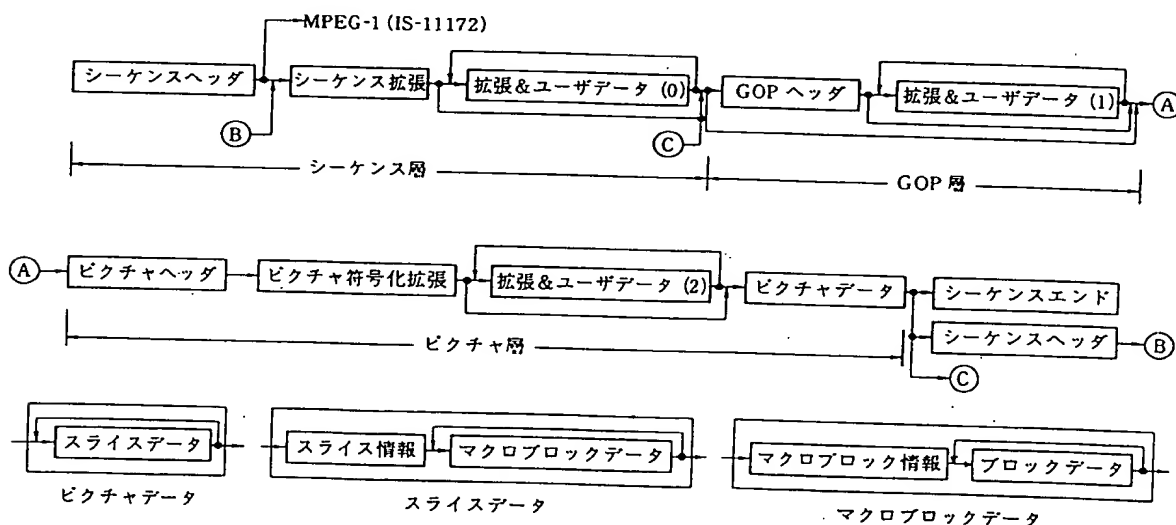


図6 MPEG ビットストリームのブロックダイアグラム

表 1(a) シーケンスヘッダー (sequence header)

コード名	ビット数	内 容
sequence header code	32	シーケンスヘッダコード
horizontal size value	12	水平方向画素数下位 12 ビット
vertical size value	12	垂直方向ライン数下位 12 ビット
aspect ratio information	4	画素アスペクト比情報
frame rate code	4	フレームレートコード
bit rate value	18	ビットレート下位 18 ビット (400 ビット単位表示)
vbv buffer size value	10	VBV バッファサイズ下位 10 ビット
intra quantiser matrix [64]	8*64	イントラ MB 用量子化マトリクス値
non intra quantiser matrix [64]	8*64	非イントラ MB 用量子化マトリクス値

表 1(b) シーケンス拡張 (sequence extension)

コード名	ビット数	内 容
profile and level indication	8	プロファイル, レベル
progressive sequence	1	シーケンス全体のプログレッシブ画像フラグ
chroma format	2	色差フォーマット
low delay	1	低遅延モード (B ピクチャなし)

表 1(c) 拡張データおよびユーザデータ 0 (Extension and user data (0))

コード名	ビット数	内 容
extension data (0)		拡張データ (0)
sequence display extension ()		シーケンス表示 ()
sequence scalable extension ()		シーケンススケーラブル拡張 ()
extension start code identifier	4	シーケンススケーラブル拡張 ID
scalable mode	2	スケーラビリティモード
layer id	4	スケーラブル階層のレイヤ ID
空間スケーラビリティの場合		
lower layer prediction horizontal size	14	予測用下位レイヤの水平サイズ
lower layer prediction vertical size	14	予測用下位レイヤの垂直サイズ
vertical subsampling factor n	5	垂直方向アップサンプル用除数
テンポラルスケーラビリティの場合		
picture mux order	3	第 1 ベースレイヤ画像前の付加レイヤ画像数
picture mux factor	3	ベースレイヤ間の付加レイヤの画像数
user data ()		ユーザデータ ()
user data	8	ユーザデータ

の情報が含まれる。

また、表 1(b) のように extension start code 後のシーケンス拡張でプロファイル、レベル、色差フォーマット、プログレッシブシーケンス (シーケンス全体

がプログレッシブ画像) が指定される他、MPEG-1 と共有しているシーケンスヘッダでの画素数やビットレートなどの上限を MPEG-2 用に拡張するための付加データが指定される。

さらに、拡張データやユーザデータ(0)では表1(c)のように、sequence display extension()において原信号のRGB変換特性や表示画像サイズなどの情報が含まれている。また、sequence scalable extension()ではスケーラビリティモード(空間、SNR、テンポラル、データパーティション)、スケーラビリティのレイヤの指定が行われる。その他、空間スケーラビリティとテンポラルスケーラビリティでのパラメータが設定される。user data(0)ではシーケンスレベルでユーザデータが自由に設定できる。

2.2 GOP 層

GOP層はGOPヘッダに始まり、1個以上のピクチャから構成される。GOP層最初の符号化画像は参照画面を用いずに独立して符号化(イントラ符号化)されるIピクチャになる。したがって、Iピクチャを用いることにより、GOPはMPEGデータからのランダムアクセスを行うポイントとして用いることができ、例えば、GOPを10~15個のピクチャで構成した場合、30フレーム/秒の画像では0.3~0.5秒毎にランダムアクセスポイントが存在する。また、GOPの長さは任意のため、シーンの切り替わりポイントでGOPを構成して、符号化効率を高めることも可能である。なお、Iピクチャは非常に大きなデータ量となり、占有バッファ量も大きくなるが、通信などのアプリケーションでは低遅延性が要求されるため、MPEG-2では図6のようにGOP層を省略することが可能で、この場合、例えば後述するスライス単位にイントラ符号化するイントラスライスを用い、数画面に分けて各スライスのイントラ符号化を巡回させることにより、Iピクチャの代用とするとともに、どの画面のバッファ占有量も平均的に低くさせることが可能である。

GOPヘッダには表1(d)のように、タイムコードの他、例えば前GOPが編集によりカットされ、GOP内の最初のIピクチャより時間的に前のBピクチャが正確に復号できないことを示すbroken linkやGOPの独立性(他のGOP内の画像に存在しない)を示すclosed gopコードがあり、これらのコードはランダムアクセス時の不当なBピクチャ表示の回避や、編集において、あるGOPをカットした時のフラグとして用いることができる。GOPヘッダに続く拡張データおよびユーザデータ(1)では、表1(e)のようにGOPレベルでのユーザデータの設定が可能である。

2.3 ピクチャ層

ピクチャ層は各画面に相当し、ピクチャヘッダ(表1(f)で画面に関する符号化条件が設定される。tem-

poral referenceは表示順番を示す画像番号で、GOP毎に0にリセットされ、picture coding typeではピクチャの符号化タイプI、P、Bが設定される。また、vbv delayでは後述するVBV(Video Buffering Verifier)の復号開始までの遅延量が設定される。

ピクチャヘッダに続くピクチャ符号化拡張(表1(g))では、f code [s] [t]で、前/後方向および水平/垂直方向の動きベクトル範囲が指定され、picture structureで3-2-3の(2)で述べるフレーム構造/フィールド構造が設定される。また、イントラマクロブロックのDC係数精度の設定とVLCタイプの選択、線形/非線形量子化スケールの選択、ジグザグ/オルタネートスキニングの選択などが行われる。

ピクチャ符号化拡張に続く拡張データおよびユーザデータ(2)では、表1(h)のようにquant matrix extension()で量子化マトリクスの設定が可能である。これにより、ピクチャ毎にイントラ/非イントラ、輝度/色差ブロックの量子化マトリクスを変更するこ

表 1(d) GOP ヘッダ (group of pictures header)

コード名	ビット数	内 容
group start code	32	GOP スタートコード
time code	25	タイムコード(時、分、秒、ピクチャ)
closed gop	1	GOP の独立性を示すグラフ
broken link	1	GOP 内Iピクチャ前のBピクチャの正当性フラグ

表 1(e) 拡張データおよびユーザデータ 1 (extension and user data(1))

コード名	ビット数	内 容
extension data(1)		拡張データ(1)
user data()		ユーザデータ()
user data	8	ユーザデータ

表 1(f) ピクチャヘッダ (picture header)

コード名	ビット数	内 容
picture stat code	32	ピクチャスタートコード
temporal reference	10	GOP 内画像の表示順序(modulo 1024)
picture coding type	3	ピクチャ符号化タイプ(I, B, P)
vbv delay	16	復号開始までの VBV 遅延量

表 1(g) ピクチャ符号化拡張 (picture coding extension)

コード名	ビット数	内 容
f code[s][t]	4	前/後方向(s), 水平/垂直(t)動きベクトル範囲
intra dc precision	2	イントラ MB の DC 係数精度
picture structure	2	ピクチャ構造(フレーム, フィールド)
top field first	1	表示フィールドの指定
frame pred frame dct	1	フレーム予測+フレーム DCT フラグ
concealment motion vectors	1	イントラ MB コンシールメント MV フラグ
q scale type	1	量子化スケールタイプ(線形, 非線形)
intra vlc format	1	イントラ MB 用 VLC タイプ
alternate scan	1	スキャニングタイプ(ジグザグ, オルタネート)
repeat first field	1	2:3 ブルダウン用フィールドリピート
chroma 420 type	1	4:2:0 の時 progressive frame と同値
progressive frame	1	プログレッシブフレームフラグ

表 1(h) 拡張データおよびユーザデータ 2 (extension and user data (2))

コード名	ビット数	内 容
extension data (2)		拡張データ (2)
quant matrix extension ()		量子化マトリクス拡張 ()
intra quantiser matrix[64]	8*64	イントラ MB 量子化マトリクス
non intra quantiser matrix[64]	8*64	非イントラ MB 量子化マトリクス
chroma intra quantiser matrix[64]	8*64	色差イントラ量子化マトリクス
chroma non intra quantiser matrix[64]	8*64	色差非イントラ量子化マトリクス
copyright extension ()		著作権拡張 ()
picture display extension ()		ピクチャ表示拡張 ()
picture spatial scalable extension ()		ピクチャ空間スケーラブル拡張 ()
spatial temporal weight code table index	2	アップサンプル用時空間重み付けテーブル
lower layer progressive frame	1	下位レイヤプログレッシブ画像フラグ
lower layer deinterlaced field select	1	下位レイヤのフィールド選択
picture temporal scable extension ()		ピクチャテンポラルスケーラブル拡張 ()
reference select code	2	参照画面の選択
forward temporal reference	10	前方向予測用下位レイヤの画像番号
backward temporal reference	10	後方向予測用下位レイヤの画像番号
user data ()		ユーザデータ ()
user data	8	ユーザデータ

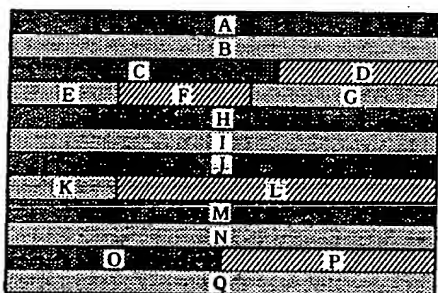
ともでき、画面の特性に応じた量子化を行うことが可能になる。また、copyright extension () では著作権番号の設定も可能になっている。picture display extension () では、表示領域の指定ができ、これにより、例えば、16:9 の HDTV を復号し、最も興味のある部分に対して、4:3 の標準テレビサイズでそ

の領域を指定することにより、標準テレビでの HDTV 画面の一部表示を行うといった、Pan-scan 機能を実現することができる。picture spatial scalable extension () や picture temporal scalable extension () では、それぞれ空間スケーラビリティやテンポラルスケーラビリティでの下位レイヤの画像

番号などの情報が設定される。なお、各ピクチャ層は図5のように1個以上のスライス層に分割されている。

2.4 スライス層

スライス層は画面内で図7のように横長の帯状の領域を示し(図のA, B, C...の領域), 画面を複数のスライスで構成することにより, あるスライス層でエラーが発生しても, 次のスライス層のスタートからの同期でエラー回復が可能となる。スライス層はMPEG-1, MPEG-2いずれの場合も1個以上のマクロブロックから構成され, ラスタスキャンオーダで左から右, 上から下に並び, その長さやスタート位置は自由で, 画面毎に変更可能である。ただし, MPEG-2では並列処理や効果的なエラー耐性を目的として1つのスライスは右方向にのみ伸び, 下方にまで伸びることはない(図7のスライスA, B, D, G等)。なお, MPEG



ピクチャ=スライスA+スライスB...+スライスQ

図7 スライス構造(MPEG-2の場合)

表1(i) ピクチャデータ (picture data)

コード名	ビット数	内 容
slice ()		スライスデータ ()

表1(j) スライス (slice)

コード名	ビット数	内 容
slice start code	32	スライススタートコード +スライス垂直位置
slice vertical position extension	3	スライス垂直位置拡張用 (>2800ライン)
priority breakpoint	7	データパーティショニング 用区分点
quantiser scale code	5	量子化スケールコード (1~31)
intra slice	1	イントラスライスフラグ
macroblock ()		マクロブロックデータ ()

-1では下方にまで伸びることができ, 1つの画面を1個のスライスのみで表すことも可能である。また, 表1(j)のようにスライスの開始の垂直位置はスライス情報中のslice start codeとslice vertical positionによって設定される。

またスライスでは, 後述する量子化スケールコード(1...31), スライス内のマクロブロックがすべてイントラ符号化であることを示すイントラスライスフラグ, データパーティショニング用区分点が設定される。

2.5 マクロブロック層

各フォーマットに対するマクロブロック構成を図8に示す。4:2:0フォーマットの場合, 図のように4つの輝度のブロックと, 空間的に対応した2つの色差ブロックの6つのブロックから構成される。

マクロブロックデータでは表1(k)のように, マクロブロックの位置や符号化モードが設定される。まず, macroblock address incrementは現在のマクロブロックの位置(macroblock address)と直前の符号化マクロブロックの位置(previous macroblock address)の差分として与えられる。また, macroblock address incrementが1より大きい場合は, 現在のマクロブロックと直前の符号化マクロブロックの間のマクロブロックはスキップマクロブロックとなり, マクロブロック内に動きベクトルやDCT係数が全くないことを示す。これにより不要なマクロブロック情報の削減を図っているが, Iピクチャについてはスキップマクロブロックはない。P, Bピクチャにおいてスライスの最初と最後のマクロブロックは必ず符号化される。

macroblock typeでは, 動き補償モード, DCT符号化モードが決定される。ピクチャタイプによって可能なモードは異なるが, 基本的には, 動きベクトルが0かどうかを示すMC/No MC選択モード(Pピクチャの場合), 動き補償予測の方向(Bピクチャの場合: 前方向, 後方向, 両方向), DCT係数の有無を示すCoded/Not Coded, イントラ符号化および量子化スケールコードの更新が設定される。

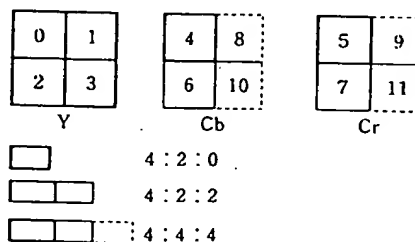


図8 マクロブロック構造

表 1(k) マクロブロックデータ (macroblock)

コード名	ビット数	内 容
macroblock escape	11	MB アドレス拡張用 (>33)
macroblock address increment	1-11	現 MB アドレスと前 MB アドレスの差
macroblock modes ()		マクロブロックモード ()
macroblock type	1-9	MB 符号化タイプ (MC, Coded 等)
spatial temporal weight code	2	アップサンプル用時空間重み付けコード
frame motion type	2	フレーム構造の動き補償タイプ
field motion type	2	フィールド構造の動き補償タイプ
dct type	1	DCT タイプ (フレーム, フィールド)
quantiser scale code	5	MB 量子化スケールコード (1~31)
motion vectors (s)		動きベクトル (s)
motion vertical field select[r][s]	1	予測に用いる参照フィールドの選択
motion vector (r, s)		動きベクトル (r, s)
motion code[r][s][t]	1-11	基本差分動きベクトル
motion residual[r][s][t]	1-8	残差ベクトル
dmvector[t]	1-2	デュアルプライム用差分ベクトル
coded block pattern ()		CBP
block (i)		ブロックデータ ()

表 1(l) ブロック (block)

コード名	ビット数	内 容
dct dc size luminance	2-9	DCT 輝度 DC 係数差分サイズ
dct dc differential	1-11	DCT 輝度 DC 係数差分値
dct dc size chrominance	2-10	DCT 色差 DC 係数差分サイズ
dct dc differential	1-11	DCT 色差 DC 係数差分値
First DCT coefficient	3-24	非イントラブロックの第 1 非零係数
Subsequent DCT coefficients	2-24	後続の DCT 係数
End of block	2 or 4	ブロック内 DCT 係数終了フラグ

動き補償予測に関しては、frame motion type および field motion type で、後述するフレーム動き補償予測やデュアルプライム動き補償予測などの動き補償予測モードが設定される。また、motion vectors () で動きベクトルが設定される。さらに空間スケラビリティでの下位レイヤのアップサンプル用の時空間重み付け係数が設定される。

DCT 符号化に関しては、coded block pattern でマクロブロック内の各 8×8 ブロックでの DCT 係数の有無が表示される他、dct type でフレーム/フィールド DCT の設定が行われる。

量子化に関しては、マクロブロックごとに量子化ステップサイズを変更できるため、画像の局所的な特性

や細かなレート制御に応じることが可能である。

2.6 ブロック層

ブロックは輝度信号または色差信号の 8 画素×8 ラインから構成され、DCT および IDCT はこの単位で行われる。

ブロックデータは量子化 DCT 係数から構成され、イントラマクロブロックの DC 成分については、隣りのブロックとの差分値に関する大きさ (dct dc size luminance, chrominance) と差分情報 (dct dc differential) が与えられ、その他の DCT 係数については非零量子化 DCT 係数までの 0 係数の長さ (ラン長) とレベルに関する情報が与えられ、EOB (End of Block) で各ブロックの DCT 係数が終了する。(中島)

3-2-3 動き補償予測符号化

1. I, B, P 符号化構造

MPEG-1, MPEG-2 いずれも、ランダムアクセス機能と高い符号化効率を得るために I (Intra coded) ピクチャ, P (Predictive coded) ピクチャ, B (Bidirectionally predictive coded) ピクチャの3つのピクチャ符号化タイプがある(図9参照)。

I ピクチャはイントラ(画面内)符号化の画面で他画面とは独立して符号化され、画面のすべてのマクロブロックがイントラ符号化になる。I ピクチャを周期的に配置することによって、ランダムアクセスやエラー回復ポイントとして用いることができる。ただし I ピクチャの頻度が高くなると総合的な符号化効率が落ちる。シーンの切り替わり点や予測効率が悪い画像については、I ピクチャの方が一般的に符号化効率が良い。

P ピクチャは前方向予測符号化画面で、時間的に過去に位置する I または P ピクチャから予測符号化を行う(図9の(i))。この画面は H.261 で用いられている予測符号化画面と同じで、前方向のみの予測を用いて符号化される。

B ピクチャは双方向予測符号化画面で、時間的に前後に位置する I または P ピクチャを用いて前方向(図9の(ii))、後方向(図9の(iii))または両方向(図9の(ii)と(iii)利用)の画面から予測符号化を行い、予測方向はマクロブロック単位に決定される。B ピクチャの導入により、物体の消失や出現のあるような領域に対しても、時間的に前後の画像を用いて的確に予測符号化を行うことができ、符号化効率を大きく向上させ

ている。ただし、その後の P ピクチャや B ピクチャの参照画面としては用いられない。また、激しい動きのある領域では、動き補償予測よりもイントラ符号化の方が符号化効率が高い場合もあるため、P ピクチャや B ピクチャではマクロブロック単位にイントラ符号化モードを用いることも可能である。

これら3種類のピクチャの組合せは自由で、符号化器が目的に合わせて設定することになる。符号化を行う順番については、時間的に後方に位置する I, P ピクチャが B ピクチャに先行して符号化されるため、例えば図9の場合、I1 P1 B1 B2 B3 P2... のようになる。なお、復号処理もこの順序で行われ、画像出力する時に元の順番に直して表示する。

一般に動きが緩やかな場合は、B ピクチャを用いることによって符号化効率は大きく向上する。逆に動きが激しい場合は、B ピクチャによる符号化効率の向上はあまり期待できない。また、通信など符号化/復号処理での遅延量が問題になるようなアプリケーションでは、B ピクチャの使用は制限される。

2. フレーム構造とフィールド構造

インタレース画像を対象としている MPEG-2 では、各画面の符号化の単位としては、フレーム構造(Frame picture)とフィールド構造(Field picture)の2つの種類があり、フレーム単位に picture structure で選択される。フレーム構造ではインタレースされた2つのフィールドを合成したフレームピクチャ単位で動き補償予測や DCT 符号化を行う(図10(a))。これに対して、フィールド構造ではインタレースされた2つのフィールドの各フィールドピクチャ単位に符号化を行う。また、P ピクチャの場合、直前に符号化された2フィールドからの動き補償予測により符号化する。ただし、前述した picture coding type(I, B, P ピクチャ)はフレーム単位で決定される(図10(b))。したがって、例えばフィールド構造で、片方のフィールドが B ピクチャの場合、他方のフィールドも B ピクチャとなる。ただし、先に符号化されたフィールドが I ピクチャの場合、他方のフィールドは I または P ピクチャのどちらも可能である。また、各フィールドはインタレース時に第1ラインがある方のフィールドをトップフィールド、他方をボトムフィールドと呼

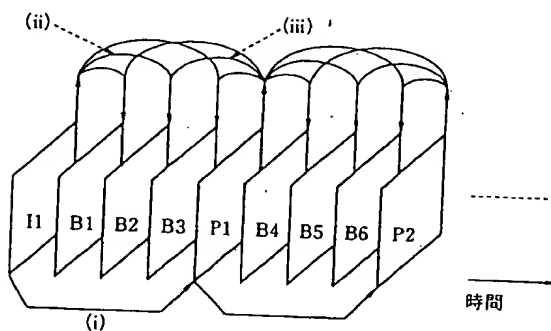
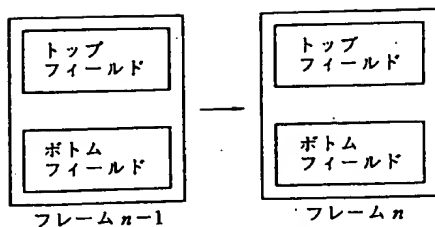
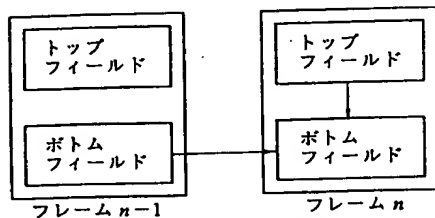


図9 IBP 予測構造



(a) フレーム構造



(b) フィールド構造

図 10 フレーム構造とフィールド構造

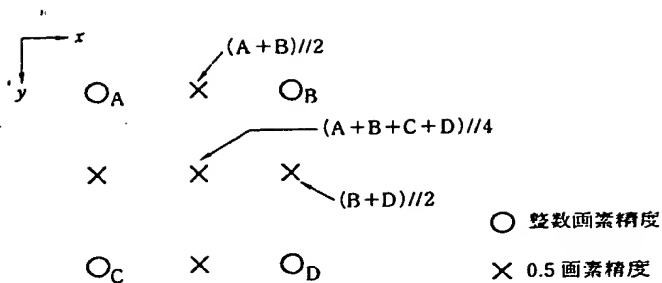


図 11 0.5 画素精度

表 2 MPEG-2 での動き補償予測モード

フレーム構造 (MV/MB)	フィールド構造 (MV/MB)
フィールド動き補償 (2)	フィールド動き補償 (1)
フレーム動き補償 (1)	16×8 動き補償 (2)
デュアルプライム (1+dmv)	デュアルプライム (1+dmv)

ぶ。

符号化効率の点からいえば、フレーム構造での符号化はインタレースされた 2 つのフィールドが保たれたまま符号化されるため、静止画像やゆっくりした動きの画像で有効である。またフィールド構造での符号化は、激しい動きや複雑な動きで、直前のフィールドからの予測の方が 2 フィールド前からの予測に比べて予測効率が高い場合に有効であるといえる。

MPEG-1 ではプログレッシブフォーマットのみを対象としており、特にこのような区別はない。

3. 動き補償予測モードと動きベクトル検出方法

MPEG-1 では、H. 261 の動き補償予測方式をベースに小数点画素精度の動き補償と両方向予測方式の導入により予測精度が大きく向上している。さらに MPEG-2 では、MPEG-1 とは違ってインタレース構造を主に対象としているため、MPEG-1 での動き補償予測方式をベースに高度な予測方式が用いられており、MPEG-2 符号化技術の中でも符号化効率の向上という点で非常に重要な位置を占めている。

予測の精度に関しては MPEG-1、MPEG-2 いずれも参照画面での各画素間を 1:1 で直線補間した 0.5 画素精度を用いており (図 11)、H. 261 でのループフィルタ (4 章参照) と同様なフィルタリング効果を持つとともに、予測精度の向上を図っている。図で // は丸め (四捨五入) 付きの除算を示す。また、動きベクトルは輝度・色差両信号に共通に用いられ、色差信号の動きベクトルは輝度信号の動きベクトルをスケールアップする。例えば 4:2:0 フォーマットの場合、動きベ

クトルを水平・垂直方向共に半分にする。

以下に MPEG-1、MPEG-2 での動き補償予測方式について述べる。

3.1 MPEG-1 での動き補償予測

DCT 符号化はブロック単位に行うが、動き補償予測については、予測効率と動きベクトル情報量のトレードオフにより、MPEG でも H. 261 と同様にマクロブロック単位に行う。MPEG-1 の画像フォーマットはプログレッシブ構造であるため、フレーム単位で輝度信号の場合 16 画素×16 ライン、色差信号の場合 8 画素×8 ラインのブロックで予測される。動きベクトルの数は P ピクチャでは、マクロブロックにつき 1 個の動きベクトルが用いられる。また、B ピクチャの場合、前方向または後方向予測ではマクロブロックにつき 1 個、双方向の場合は 2 個の動きベクトルが用いられる。なお、双方向動き補償予測は画面間の距離にかかわらず、1:1 で前方向と後方向の予測画像を直線補間した画像を用いる。

3.2 MPEG-2 での動き補償予測

MPEG-2 での動き補償予測モードを表 2 に示す。表で MV/MB は前方向予測の場合について 1 マクロブロック当たりの動きベクトルの数を示す。

(1) フレーム構造での動き補償予測

まず、フレーム構造では、(a) フレーム動き補償予測、(b) フィールド動き補償予測、および (c) デュアルプライム予測という 3 つの動き補償予測方式が用いられる。

(a) フレーム動き補償予測は、インタレースされた 2 つのフィールドが合成されたフレームで動き補償

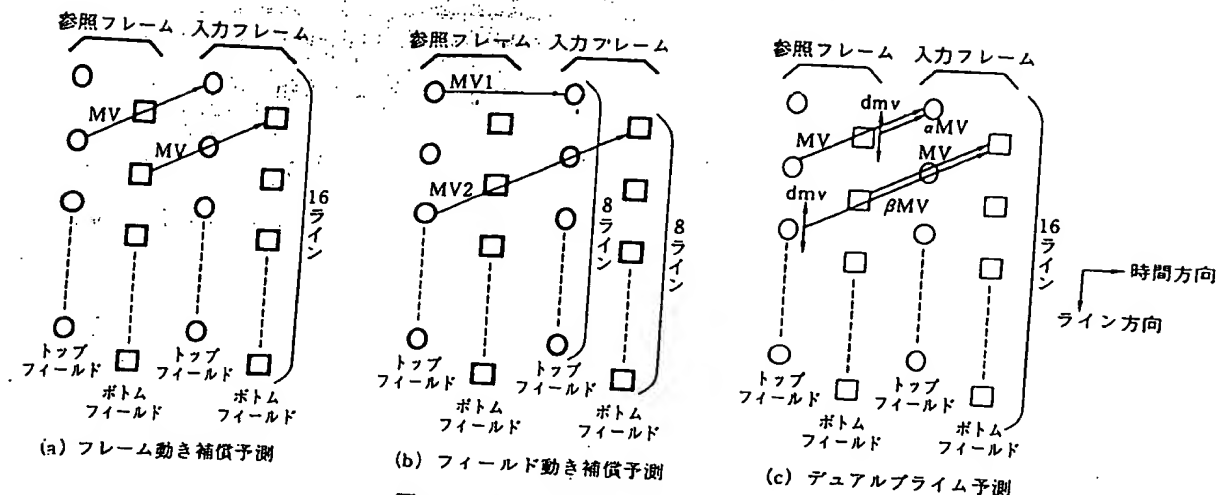


図 12 フレーム構造での動き補償予測

予測を行うもので、輝度信号はインタレースされた 16 画素×16 ラインブロックごとに予測される(図 12 (a))。図は 1 フレーム離れた参照フレームから前方の動き補償予測を行う例で、横軸が時間、縦軸がライン方向を示す。フレーム動き補償予測は比較的ゆっくりとした動きで、フレーム内での相関が高いまま等速度で動いている場合に有効な予測方式である。

フレーム動き補償予測は、インタレース構造である点を除けば MPEG-1 と同じ予測方式で、動きベクトルの数についても MPEG-1 と同様である。

(b) フィールド動き補償予測は、フィールドごとに動き補償を行うもので、図 12(b) のようにトップフィールドには MV1、ボトムフィールドには MV2 を用いて予測を行っている。また、参照フィールドはトップフィールドでもボトムフィールドでもよく、マクロブロックデータ中の motion vertical field select フラグで設定される。図 12(b) ではトップフィールド、ボトムフィールドいずれにもトップフィールドが参照フィールドとして用いられている。フィールド動き補償予測では、マクロブロック内の各フィールド別に予測されるため、輝度信号の場合 16 画素×8 ラインのフィールドブロック単位で予測される。

動きベクトルの数に関しては、P ピクチャや B ピクチャの前方向や後方向予測ではマクロブロックにつき 2 個必要になる。また B ピクチャの双方向予測では 4 個の動きベクトルが必要になる。このため、フィールド動き補償予測では、フィールド別に予測して、局所的な動きや加速度的な動きに対して予測効率を高めることが可能な反面、動きベクトル数はフレーム動き補償に比べ 2 倍必要となるため、総合的な符号化効率は低下する可能性がある。

(c) デュアルプライム予測は、各フィールドブ

ックに対して 2 つの参照フィールドブロックの平均値により予測を行うものである(図 12(c))。図でトップフィールドブロック(16 画素×8 ライン)は、動きベクトル MV で指定される参照画面のトップフィールドブロック(16×8)と $\alpha MV + dmv$ で指定されるボトムフィールドブロック(16×8)の平均値で予測される。またボトムフィールドブロックに関しては、MV で指定される参照画面のボトムフィールドブロックと $\beta MV + dmv$ で指定されるトップフィールドブロックの平均値で予測する。ここで、 α, β はフィールド間の距離比で動きベクトルをスケーリングするためのスケーリングファクタで、 $\alpha = 1/2, \beta = 3/2$ である。また、 dmv は微調整用動きベクトルで、水平・垂直方向に ± 0.5 画素の範囲で調整を行う。したがって、実際に伝送されるベクトル数は基本となる MV と微調整ベクトル dmv のみになる。

デュアルプライム予測の大きな特徴は、1 つのフィールドブロックに 2 つのフィールドブロックを用いた時間方向フィルタリング作用と微調整ベクトルによる空間解像度の向上によって予測効率を大きく向上させたことである。また、伝送に必要な動きベクトルがマクロブロック毎に 1 個のベクトルと、符号化情報量の小さい微調整ベクトル 1 個ですむため、符号化情報量を含め総合的に符号化効率を高めている。

ハードウェアの構造上は、2 つのブロックから予測を行う点で、B ピクチャの双方向予測と同程度の処理速度を要求される。したがって、デュアルプライムを双方向予測に用いた場合、その他の予測方式での双方向予測の 2 倍の処理量が要求される。この場合、特にメモリーバンド幅の負担が非常に大きくなるため、デュアルプライム予測は、予測される画像と参照画像との間に B ピクチャのない P ピクチャ(例えば SP

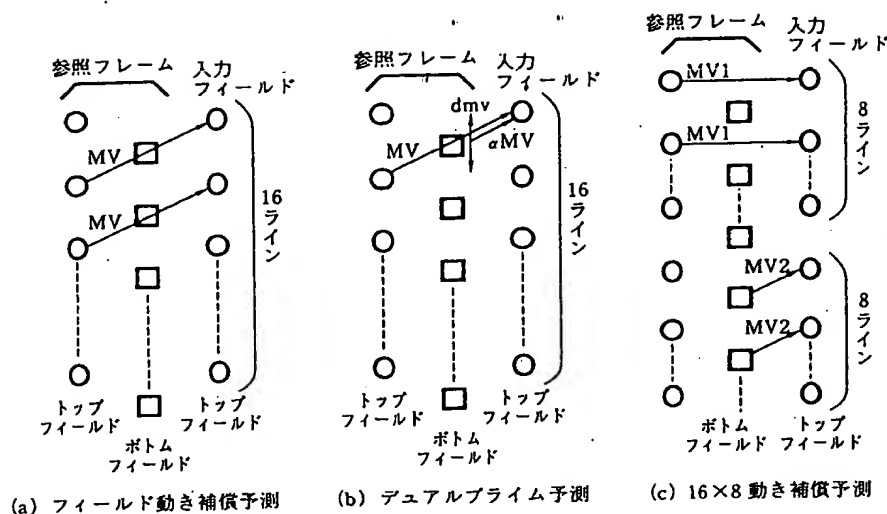


図 13 フィールド構造での動き補償予測

@ML)でしか用いることはできない。

これまでの評価では、フレーム/フィールド/デュアルプライム予測の性能は、B ピクチャが入った場合のフレーム/フィールド予測方式とほぼ同程度の性能が得られることが、いくつかの評価画像において確認されている⁷⁾。

(2) フィールド構造での動き補償予測

フィールド構造では、フィールド動き補償予測、デュアルプライム予測、16×8 動き補償予測の3つの予測方法がある。フィールド構造の場合のマクロブロックは、トップまたはボトムフィールドいずれかで16画素×16ライン(輝度信号)のフィールドブロックが構成されている。このため、フィールド動き補償予測では図13(a)のように16画素×16ライン単位で予測を行う。なお、動きベクトル数はフレーム構造のフレーム動き補償予測と同様に、Bピクチャの双方向予測でマクロブロックにつき2個で、それ以外は1個/マクロブロックになる。

フィールド構造でのデュアルプライム予測については、予測するブロックの大きさが16画素×16ラインのフィールドブロック単位であることを除けば、フレーム構造でのデュアルプライム予測と同じで、図13(b)のように2つの参照フィールドブロックからの予測の平均値を用いて動き補償予測を行う。

16×8 動き補償予測は、16画素×16ラインのフィールドマクロブロックを図13(c)のように上部8ラインの16×8ブロックと下部8ラインの16×8ブロックに分割して予測を行う(4:2:0フォーマットの場合、色差ブロックは2つの8×4ブロックに分割される)。図のように上半分のフィールドブロックには動きベクトルMV1を用い、下半分のブロックには

MV2を用いて動き補償予測を行う。さらに、図のように参照するフィールドが上部ブロックと下部ブロックで異なることも可能で、より細かな動きに対して予測効率を高めることが可能である。しかしながら、動きベクトル数がフィールド動き補償予測の2倍必要のため、総合的な符号化効率としては大きくは向上しない恐れがある。

3.3 動きベクトル予測方法

動きベクトル予測は符号化側の処理であるため、その方法はMPEGの標準の範囲外であるが、標準化の過程で用いられたテストモデルTM4⁴⁾では以下のような方法が利用された。ここではMPEG-2でのフレーム構造でのフレーム/フィールド動き補償予測の場合について説明する。

まず、符号化される前の原画面を用いて、探索範囲内での輝度信号16画素×16ラインの予測誤差を整数精度でブロックマッチング法を用いて求める。例えば、水平方向の探索範囲を+/-15画素/フレームとした場合、図9ではPピクチャは参照画面から4フレーム離れているため、+/-60画素の範囲を探索する。予測誤差については、参照画像と入力画像のブロックでの差分絶対値の累積和AEを用いる。フィールドブロック毎に動き予測誤差(AE of Top/Bottom Field)を求めることにより、フレーム動き予測誤差(AE of Frame)は

$$\begin{aligned} (\text{AE of Frame}) &= (\text{AE of Top Field}) \\ &+ (\text{AE of Bottom Field}) \end{aligned} \quad (1)$$

で求めることができる。次に、探索範囲内のすべての位置においてトップフィールド、ボトムフィールド、フレームの各予測誤差を求め、それぞれが最小となる

位置とその位置での予測誤差を求め、整数精度での探索結果とする。

次に符号化後の局所復号画面において、整数精度で求めた各予測の最小予測誤差位置を中心に、上下左右に0.5画素ずらした8つの位置での予測誤差を求め、予測誤差が最小となる位置と誤差値を求める。

Bピクチャの場合、前方向、後方向それぞれ上記の方法で動きを求める。双方向予測については、前方向と後方向でそれぞれ予測誤差が最小となる位置で、1:1で線形内挿した参照画像を用いて予測誤差を求める。

4. 動きベクトルの符号化

画像の近傍領域における動き量は相関が高いという特性を利用して、動きベクトルの符号化については、前に符号化したマクロブロックの動きベクトル量を予測ベクトルとして、その予測ベクトルとの差分ベクトルを符号化している。

MPEG-1では前/後方向動きベクトルで、最大2つの動きベクトルがマクロブロックに存在するため、予測ベクトルもそれに合わせて用意される。またMPEG-2では、例えば、フレーム構造のフィールド動き補償では各フィールドに1個の動きベクトルが必要となるため、双方向予測の場合、3-2-3節の3.2(1)で述べたように、前後方向の動きベクトルと合わせて合計4つの動きベクトルが1つのマクロブロックに必要となる。

動きベクトルは $\text{vector}'[r][s][t]$ 、予測ベクトルは $\text{PMV}[r][s][t]$ と表され、差分ベクトルを delta とすると、動きベクトル $\text{vector}'[r][s][t]$ は

$$\text{vector}'[r][s][t] = \text{PMV}[r][s][t] + \text{delta};$$

(2)

と表される。ただし、 $r, s, t = 0/1$ で、 r はマクロブロックの第1/第2動きベクトル、 s は前方/後方動きベクトル、 t は水平/垂直成分を示す。フレーム構造のフィールド動き補償予測では、第1/第2動きベクトルはトップ/ボトムフィールド動きベクトルに対応する。また、(2)式のベクトルは0.5画素も表現するために、2倍した整数値が用いられる。さらに、予測ベクトルの $\text{PMV}[r][s][t]$ はスライスの先頭、イントラマクロブロックで、0にリセットされる。

差分ベクトル delta は、基本差分ベクトルをスケールファクタによりスケールし、残差を加えて表現する。基本差分ベクトルを motion code 、スケールファクタを f 、残差ベクトルを motion residual とすると、 delta は以下の式で表される。

$$\text{delta} = \text{Sign}(\text{motion code}) * \{ ((\text{Abs}(\text{motion code}) - 1) * f) + \text{motion residual} + 1 \} \quad (3)$$

ここで、基本差分ベクトル motion code は-16から+16までの整数で、可変長符号化コード(1~11ビット)としてマクロブロックに与えられる。またスケールファクタ f は、動き補償範囲を決定するスケールファクタコード $f \text{ code}[s][t]$ (ピクチャ毎に設定される)を用いて、

$$f = 1 \ll (f \text{ code}[s][t] - 1) \quad (4)$$

と表される。また、残差ベクトル motion residual はマクロブロックに与えられ、 $(f \text{ code} - 1)$ のビット長を持つ。

例えば、スケールファクタコード $f \text{ code} = 8$ 、基本差分ベクトル $\text{motion code} = -16$ 、残差ベクトル $\text{motion residual} = 127$ と与えられた場合、(3)(4)式を用いて、差分ベクトル $\text{delta} = -2048$ となる。

このように差分ベクトル delta は、スケールファクタコード、33個の動きベクトルを持つ motion code テーブルと残差ベクトルによって、広い範囲を効率的に符号化することが可能である。

また、各スケールファクタ f に対する動きベクトルの最大、最小および範囲はそれぞれ以下のように設定される。

$$\begin{aligned} \text{high} &= (16 * f) - 1, \quad \text{low} = ((-16) * f), \\ \text{range} &= (32 * f) \end{aligned} \quad (5)$$

例えば $f = 2$ の場合、 $\text{high} = 31$ 、 $\text{low} = -32$ となり、動きベクトルの範囲は、0.5画素精度で-16.0から+15.5までとなる。なお、差分ベクトル delta の範囲は動きベクトルの2倍の範囲が必要になるが、補数表現を用いて delta も動きベクトルと同じ範囲で表現される。例えば、表3のように、動きベクトルの範囲が-128から127で、 $\text{PMV} = -128$ の場合、動きベクトル vector' が-128から-1までの範囲は差分ベクトル $\text{delta} = \text{vector}' - \text{PMV}$ は0から127になるが、 vector' が0から127の間は、差分ベクトル $(\text{vector}' - \text{PMV})$ は128から255になる。したがってこれらの範囲は、本来、動きベクトルの範囲外になる-128から-1を差分ベクトル delta として用いる。

表3 差分ベクトル delta の補数表現例

vector'	-128	-1	0	127
PMV	-128					
$\text{vector}' - \text{PMV}$	0	127	128	255
delta	0	127	-128	-1

MPEG-2のMP@MLでは、フレーム構造のフレーム動き補償予測の場合、動きベクトルの範囲は垂直方向で $[-128.0, +127.5]$ 、水平方向で $[-1024.0, 1023.5]$ となり(f_code がそれぞれ8, 5)、MPEG-1の基本設定(Constraint parameter)時の $[-64.0, 63.5]$ に比べて飛躍的に大きな範囲が利用可能である。なお、MPEG-2での最大動きベクトル範囲は $[-2048.0, 2047.5]$ で、例えば、MP@HLの水平方向ではこの範囲が用いられる。また、フィールド構造については、動きベクトル範囲はフレーム構造の場合の半分である。

5. マクロブロックモードの決定方法

TM4では以下のようにして各種モードの決定を行っている(フレーム構造の場合)。

5.1 Pピクチャの動き補償予測

フレーム構造のPピクチャでは、フレーム/フィールド/デュアルプライムの3つの動き補償予測が可能であるが、それぞれの予測方式での輝度ブロック(16画素 \times 16ライン)の二乗予測誤差を比較して、最小となる方式を選択する。また、動きベクトルが水平・垂直ともに0で現画像と全く同位置で予測を行う場合、NoMCモードとし、動きベクトルの水平・垂直成分の

いずれかが0でない場合、MCモードとする。

5.2 Bピクチャの動き予測

フレーム構造のBピクチャの動き補償予測については、前方向/後方向/双方向モードとフレーム/フィールド動き予測モードにより合計6種類の予測モードが可能であるが、すべての場合について二乗予測誤差を求め、最小となるモードを選択する。もし誤差値が同一の場合は、基本的に符号化情報量の低い方を優先し、

フレーム予測 > フィールド予測

前方向 > 後方向 > 双方向予測

の順としている。

5.3 イントラ/インター符号化モード

Pピクチャ、Bピクチャいずれの場合も動き予測(インター)符号化かイントラ符号化かを以下のようにして選択する。

まず、現画像の輝度ブロック16 \times 16の平均二乗偏差VARORを求める。次に、動き予測時の平均二乗予測誤差VARを求める。その結果、VARが小さい場合は動き予測モードを優先して、VAROR < VARかつVAR > 64であればイントラ符号化、それ以外ではインター符号化とする。(中島)

3-2-4 DCT符号化, 量子化, 可変長符号化

1. フレーム/フィールド DCT 符号化

各マクロブロックの予測誤差信号(イントラマクロブロックでは入力信号)は、ブロックに分解され、各ブロックに対して8 \times 8 DCT符号化が行われる。

ここで、MPEG-2のフレーム構造においては、2種類のDCT符号化モードをマクロブロック単位に切

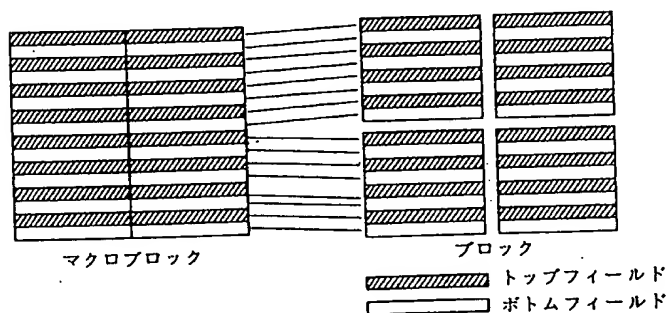


図 14(a) フレーム DCT モードにおけるマクロブロック輝度信号とブロックの関係

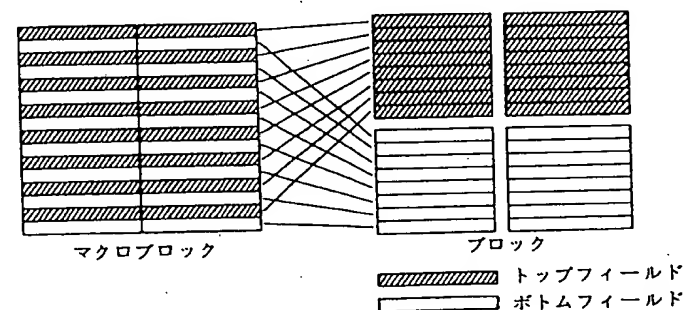


図 14(b) フィールド DCT モードにおけるマクロブロック輝度信号とブロックの関係

ては、図 14(b)のように各ブロックがフィールドで構成されるように分解される。

2つの DCT 符号化モードを適応的に切り替えることにより、動き補償におけるフレーム/フィールド適応処理と併せて、インタレース信号に対する符号化効率が向上する。

なお、4:2:0 フォーマットにおける色差信号については、8×4 DCT の適用を避けるため、常にフレームに対し 8×8 DCT を施す。

また、フィールド構造においては、マクロブロックが 1 フィールドの信号のみで構成されるため、輝度信号と色差信号の両者について、常にフィールド DCT を施す。

どちらの DCT 符号化モードを選択するかは決定方法は標準化の範囲外であるが、例えば TM4⁹⁾においては以下の決定方法が採られている。まず、各マクロブロックについて、(7)式に示すフレーム内で垂直方向に隣接した画素間差分値の二乗和 Var1 と、(8)式に示すフィールド内で垂直方向に隣接した画素間差分値の二乗和 Var2 を求める。次に、Var1 が Var2 に比較して大きく(6)式が真となった場合は、フィールド毎に DCT を施した方が変換効率が高くなると判断し、フィールド DCT を選択する。

$$\text{Var1} > \text{Var2} \quad (6)$$

$$\text{Var1} = \sum_{u=0}^{15} \sum_{v=0}^6 \{ (X[2v][u] - X[2v+1][u])^2 + (X[2v+1][u] - X[2v+2][u])^2 \} \quad (7)$$

$$\text{Var2} = \sum_{u=0}^{15} \sum_{v=0}^6 \{ (X[2v][u] - X[2v+2][u])^2 + (X[2v+1][u] - X[2v+3][u])^2 \} \quad (8)$$

$X[v][u]$: マクロブロック輝度信号の第 (u, v) 要素

また、デコードおよびエンコード局部復号化部では逆変換(IDCT)が実行される。ここで IDCT は実数演算で定義されており、しかも、その演算方式には様々な高速演算方式が存在する。したがって、IDCT の演算方式や演算精度の違いにより、エンコード局部復号化部とデコードにおける IDCT 演算結果が必ずしも同一とならず、ミスマッチが生じる。さらに、このミスマッチは予測の繰り返しによって蓄積される。

そこで、IDCT 演算精度については、IEEE 規格⁹⁾に示されている演算精度を満足することを規定している。IEEE 規格では、あらかじめ決められた方法で発生される一連のテストデータに対して、評価の対象となる IDCT 演算結果と、十分に高い演算精度で演算された IDCT 演算結果とを比較して、平均二乗誤差

等の統計量を算出し、その値が定められた上限値以下である場合に、評価対象となる IDCT が規定された演算精度を満足すると定めている。

本規定は、特に H. 261 規格を考慮して定められたものであり、H. 261 規格における非イントラマクロブロックの繰り返し回数(132 回)においても、十分な画質が保たれるように上記統計量の上限値が定められている。したがって、MPEG ビデオ規格においても、P ピクチャにおける非イントラマクロブロックの繰り返しについては、その回数を 132 回までに制限することを規定している。なお、IDCT 演算に関する本規定は、MPEG ビデオ規格の中で唯一演算誤差を許す部分である。

さらに MPEG では、後述するように IDCT 演算誤差に伴うミスマッチを最小限に留める対策が組み込まれている。これに関しては次節で説明する。

なお、MPEG-1 ではコンフォーマンス規格⁹⁾において、その精度を明記すれば IEEE 規格より精度の低い IDCT 演算も許容している。これは、特にソフトウェアデコータでの負担軽減を考慮して規定されたものである。一方、MPEG-2 では高画質を維持するため、精度の低下を許容していない。

2. 量子化処理と IDCT ミスマッチ対策

MPEG ビデオ規格において、量子化は逆量子化の規定に含まれる幾つかのパラメータを変化させることにより、高画質化や視覚特性を反映した符号化を目的として、その自由度の範囲内で制御可能である。

イントラマクロブロックの DC 係数量子化については、ピクチャ単位にその量子化精度を指定することが可能であり、その他の係数については、ピクチャ単位で指定可能な量子化マトリクスの各要素に、マクロブロック単位で指定可能な量子化スケールを乗じた値により各係数の量子化精度を制御できる。

MPEG-2 においては、高画質化を図るために MPEG-1 の量子化処理に対して幾つかの拡張がなされた。新たに導入された主な点は以下の通りである。

- ・ DC 係数量子化精度の向上
- ・ 非線形量子化スケールの導入
- ・ 高画質対応 IDCT ミスマッチ対策の導入

2.1 イントラマクロブロックの DC 係数逆量子化
イントラマクロブロックの DC 係数は、(9)式に従って逆量子化される。

$$F^{-1}[0][0] = \text{intra dc mult} \times \text{QF}[0][0] \quad (9)$$

$F^{-1}[0][0]$: DC 係数値の量子化代表値

$\text{QF}[0][0]$: DC 係数値の量子化代表値レベル番号

intra dc mult は、DC 係数量子化精度を指定するためにピクチャ単位で設定可能なパラメータである。intra dc precision に従い、表 4 に示す関係で決まる値である。

MPEG-1 においては、intra dc precision が 0 に対応する精度(8 ビット精度相当)のみであったが、輝度レベルがゆるやかに変化する画像を高画質で符号化するためには、この精度では不充分であった。MPEG-2 においては、本パラメータを用いて 8～11 ビット精度相当の DC 係数量子化精度を指定することが可能となった。ただし、4:2:2 フォーマットの使用が可能であり、画質への要求レベルが特に高い High プロファイル以外では、最高 10 ビット精度で充分とし、8～10 ビット精度に制限している。

2.2 その他の係数の逆量子化

その他の係数は、(10)式に従って逆量子化される。

$$F''[v][u] = ((2 \times QF[v][u] + k) \times W[w][v][u] \times \text{quantiser scale}) / 32 \quad (10)$$

$F''[v][u]$: 第 (u, v) 係数値の量子化代表値

$QF[v][u]$: 第 (u, v) 係数値の量子化代表値レベル番号

$$k = \begin{cases} 0 & \text{イントラマクロブロック} \\ \text{Sign}(QF[v][u]) & \text{非イントラマクロブロック} \end{cases} \quad (11)$$

ここで $W[w][v][u]$ および quantiser scale は、それぞれ量子化マトリクスおよび量子化スケールであり、これらのパラメータによって量子化特性が制御される。

またパラメータ k は、非イントラマクロブロックにおいて、その値が $QF[v][u]$ の符号に応じて 1, 0 または -1 となる。例えば、 $QF[v][u]$ が -2, -1, 0, 1, 2 に対し、 $F''[v][u]$ は、それぞれ $-5k, -3k, 0, 3k, 5k$ (k は定数) となり、零近傍にデッドゾーンを設けている。

(1) 量子化マトリクス

量子化マトリクス $W[w][v][u]$ は、量子化特性を

表 4 intra dc precision とビット精度、逆量子化係数、DC 予測リセット値との関係

intra dc precision	ビット精度	逆量子化係数 (intra dc mult)	DC 予測 リセット値
0	8	8	128
1	9	4	256
2	10	2	512
3	11	1	1024

視覚特性に合致させるためのパラメータで、ピクチャ単位に設定可能である。MPEG-1 および MPEG-2 の 4:2:0 フォーマットの場合、2 種類の量子化マトリクス(イントラと非イントラの各マクロブロック用)が設定でき、4:2:2 または 4:4:4 フォーマットの場合、さらに輝度信号と色差信号で独立の合計 4 種類の量子化マトリクスが設定できる。 $W[w][v][u]$ における w (0～3) は、4 種類の量子化マトリクスのうちの 1 つを指定する。

表 5 に、量子化マトリクスのデフォルト値を示す。量子化マトリクスはユーザがピクチャ単位で設定可能であるが、設定されていない場合には、このデフォルト値が用いられる。デフォルト値はイントラマクロブロックのみ重みを有している。なお、MPEG-2 の TM において使用された非イントラマクロブロックに対する量子化マトリクスを表 6 に示す。これはデフォルト値とは異なり、非イントラマクロブロックについても重みを有した特性となっている。

表 5(a) イントラマクロブロック量子化マトリクス $W[0][v][u]$ のデフォルト値

	u							
	0	1	2	3	4	5	6	7
0	8	16	19	22	26	27	29	34
1	16	16	22	24	27	29	34	37
2	19	22	26	27	29	34	34	38
3	22	22	26	27	29	34	37	40
4	22	26	27	29	32	35	40	48
5	26	27	29	32	35	40	48	58
6	26	27	29	34	38	46	56	69
7	27	29	35	38	46	56	69	83

表 5(b) 非イントラマクロブロック量子化マトリクス $W[1][v][u]$ のデフォルト値

	u							
	0	1	2	3	4	5	6	7
0	16	16	16	16	16	16	16	16
1	16	16	16	16	16	16	16	16
2	16	16	16	16	16	16	16	16
3	16	16	16	16	16	16	16	16
4	16	16	16	16	16	16	16	16
5	16	16	16	16	16	16	16	16
6	16	16	16	16	16	16	16	16
7	16	16	16	16	16	16	16	16

表 6 テストモデルにおける非イントラマクロブロック量子化マトリックス $W[1][v][u]$ のデフォルト値

	u							
	0	1	2	3	4	5	6	7
0	16	17	18	19	20	21	22	23
1	17	18	19	20	21	22	23	24
2	18	19	20	21	22	23	24	25
3	19	20	21	22	23	24	26	27
4	20	21	22	23	25	26	27	28
5	21	22	23	24	26	27	28	30
6	22	23	24	26	27	28	30	31
7	23	24	25	27	28	30	31	33

v

視覚的に最適な量子化マトリックスは、入力画像に依存するため、画像に応じて切り替えることにより画質の向上が見込める。

(2) 量子化スケール

量子化スケール(quantiser scale)は、量子化特性のスケールリングを行うことにより発生符号量を制御するためのパラメータであり、ピクチャ単位で設定されるパラメータ q scale type とマクロブロック単位で設定される量子化スケールコード(quantiser scale code)により決定される。図 15 にこれらの関係を示す。

q scale type が 0 の場合は線形量子化となり、MPEG-1 と同様に quantiser scale code(1~31)の 2 倍の値が quantiser scale(2~62)に設定される。

q scale type が 1 の場合は非線形量子化となり、quantiser scale code(1~31)は、小さい量子化スケールコードではより細かく、大きな量子化スケールコードではより粗くスケールリングすることにより、前者の場合に比較して広い範囲の quantiser scale(1~112)に変換される。MPEG-2 においては、本モードを用いることにより、特に高レートで符号化する場合や、極めて複雑な画像を符号化する場合において、MPEG-1 に比較してより安定な符号量制御が可能となる。

2.3 IDCT ミスマッチ対策

DCT および IDCT 演算は実数演算で定義されているが、その間で量子化処理が入るため、IDCT 演算結果は DCT 入力値とは必ずしも一致せず、しかも整数になるとも限らない。したがって、演算結果の小数点以下の値が 0.5 である場合、個々の IDCT 演算方式の演算精度の違いによって、丸めの方向を一致させることができず、ミスマッチが生じる。

IDCT ミスマッチ対策の目的は、逆量子化後の係数

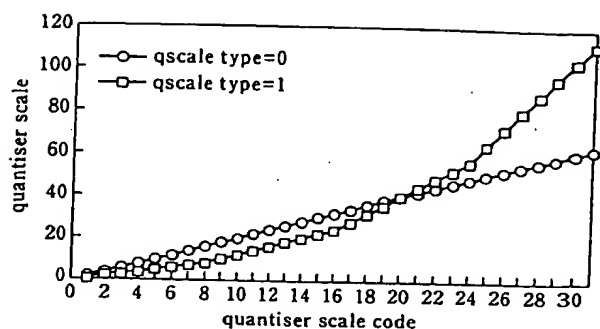


図 15 量子化スケールコード(quantiser scale code)と量子化スケール(quantiser scale)の関係

値を微小に変化させて、誤差を伴わない IDCT 実数演算結果の小数点以下の値が 0.5 となる確率をできる限り減らすことにある。

(1) MPEG-1 における IDCT ミスマッチ対策

MPEG-1 における IDCT ミスマッチ対策および係数範囲制限処理は、(12)、(13)式で示されるように、逆量子化後の各係数値をそれぞれ零の場合を除き奇数化した後、係数値を $-2048 \sim 2047$ に制限する。

$$F'[u][v] = \begin{cases} F''[v][u] - \text{sign}(F''[v][u]) & F''[v][u] : \text{偶数} \\ F''[v][u] & F''[v][u] : \text{奇数} \end{cases} \quad (12)$$

$$F[v][u] = \begin{cases} 2047 & F'[v][u] > 2047 \\ F'[v][u] & -2048 \leq F'[v][u] \leq 2047 \\ -2048 & F'[v][u] < -2048 \end{cases} \quad (13)$$

$F[v][u]$: IDCT ミスマッチ対策、係数範囲制限処理後の第 (u, v) 係数値

この対策は、IDCT ミスマッチを発生させる様々な場合のうち、特に低レートで多く発生するブロック内の単一の係数のみが非零である場合におけるミスマッチの回避を行うものである。ブロック内の逆量子化後の係数値のうち $F''[0][0]$, $F''[0][4]$, $F''[4][0]$, または $F''[4][4]$ のみが非零でその値が $8m+4$ (m は整数) の場合、IDCT 出力値 $f[y][x]$ は $\pm(m+1/2)$ (符号は x, y に依存) となりミスマッチが発生する。したがって、非零係数を奇数化することにより、ミスマッチが回避できる。

(2) MPEG-2 における IDCT ミスマッチ対策

MPEG-1 における IDCT ミスマッチ対策は、すべての非零偶数係数について値を変化させてしまうため、高レート符号化時において S/N を向上させる妨げとなる。また、複数の係数が非零となる場合に発生するミスマッチに対しては対応していない。そこで

MPEG-2においては、MPEG-1に比較し係数値の変化が少なく、かつ複数の係数が非零となる場合にも有効な新しいミスマッチ対策が導入された。

MPEG-2においては、(14)～(17)式に示すように、まず逆量子化後の各係数値を-2048から2047に制限し、各ブロックにおける係数値の総和が偶数の時、第(7, 7)係数値を、2の補数表現としてLSBを反転させる。

$$F'[v][u] = \begin{cases} 2047 & F''[v][u] > 2047 \\ F''[v][u] & -2048 \leq F''[v][u] \leq 2047 \\ -2048 & F''[v][u] < -2048 \end{cases} \quad (14)$$

$$\text{sum} = \sum_{v=0}^{u \leq 8} \sum_{u=0}^{u \leq 8} F'[v][u] \quad (15)$$

$$F'[v][u] = F'[v][u] : u=7 \text{ かつ } v=7 \text{ を除く} \\ \text{すべての } u, v \quad (16)$$

$$F[7][7] = \begin{cases} F'[7][7] & \text{sum : 奇数} \\ \begin{cases} F'[7][7]-1 & F'[7][7] : \text{奇数} \\ F'[7][7]+1 & F'[7][7] : \text{偶数} \end{cases} & \text{sum : 偶数} \end{cases} \quad (17)$$

この対策は、ブロック内の単一の係数のみが非零である場合に対するミスマッチ回避に加え、さらに以下に示す複数の係数が非零である場合に発生するミスマッチの回避も可能である¹¹⁾¹²⁾。

- ・ブロック内の逆量子化後の係数値のうち $F''[0][0]$, $F''[0][4]$, $F''[4][0]$ および $F''[4][4]$ の4個の係数のうち2個以上が非零で、 $F''[0][0] \pm F''[0][4] \pm F''[4][0] \pm F''[4][4]$ が $8m+4$ (m は整数) となり、ある x, y に対する IDCT 出力値 $f[y][x]$ が $(m+1/2)$ となる場合
- ・ブロック内の逆量子化後の係数値のうち、ある決まった2個の係数(例えば $F''[1][3]$ と $F''[3][1]$) だけが非零でかつ等しい値で、さらにその値が $8m+4$ (m は整数) となり、ある x, y に対する IDCT 出力値 $f[y][x]$ が $(m+1/2)$ となる場合

MPEG-1のIDCTミスマッチ対策が、すべての係数を奇数化するのに対し、本対策は第(7, 7)係数値のみを変えるだけであり、高レートでの符号化時にS/Nを落とさずにミスマッチの低減を実現できる。

例えば、いくつかのテスト画像を用いて8Mb/sで符号化した場合、MPEG-1のミスマッチ対策では700～3000個程度のブロックでミスマッチが発生したが、本対策では1～20個程度に収まったという報告がある¹¹⁾。

3. 適応スキニングと可変長符号化

3.1 イントラマクロブロックDC係数に対する処理

イントラマクロブロックにおける各ブロックのDC係数は、以下に示すDC予測によりブロック間相関を低減した後、予測誤差を可変長符号化する。

(1) DC予測

DC予測は、各ブロックのDC係数についてY, Cr, Cbそれぞれ独立に、図8に示されるマクロブロック内のブロック符号化順に従った1ブロック前のDC係数を予測値とし、予測値との差分値をDC予測誤差とする。

ただし、以下の条件においては、予測値をリセットする。

- ・スライス先頭
- ・非イントラマクロブロック
- ・スキップマクロブロック

リセット値は、DC係数量子化精度を指定するパラメータ intra dc precision に従い表4に示す値となる。

(2) DC予測誤差VLC

DC予測誤差は、まずその大きさ(ビット数)を“サイズ”(dct dc size)として可変長符号化し、予測誤差そのものの値は符号長 dct dc size の固定長符号(dct dc differential)で符号化する。表7にDC予測誤差のサイズに対する可変長符号を示す。例えばDC予測誤差が255の場合、サイズ dct dc size は8となり、dct dc differential は8ビットで表せる。

表7 DC差分値のサイズ(dct dc size)に対する可変長符号

予測誤差	dct dc size	VLC (luminance)	VLC (chrominance)
0	0	100	00
1	1	00	01
2-3	2	01	10
4-7	3	101	110
8-15	4	110	1110
16-31	5	1110	1111 0
32-63	6	1111 0	1111 10
64-127	7	1111 10	1111 110
127-255	8	1111 110	1111 1110
256-511	9	1111 1110	1111 1111 0
512-1023	10	1111 1111 0	1111 1111 10
1024-2047	11	1111 1111 1	1111 1111 11

3.2 それ以外の係数に対する処理

それ以外の係数は、適応スキニングにより2次元係数を1次元に並べ換え、連続する零係数の数(run)と非零係数(level)のペアを事象とした2次元可変長符号化を行う。

(1) 適応スキニング

2次元係数を1次元に並べ換える際のスキャン順序は、ピクチャ単位で設定可能なパラメータ alternate scan により、2つのスキャン順序から選択できる。

alternate scan が 0 の場合、スキャン順序は MPEG-1 と同一である(ジグザグスキャン)。alternate scan が 1 の場合は、インタレース信号に有効な縦方向スキャンに近いスキャン順序となる(4章を参照)。

(2) 2次元 VLC

run と level のペアに対する符号は、その事象に応じて次の3種類のいずれかの符号となる。

(a) 通常の符号

比較的発生頻度の高い run, level のペアに対しては、定められた符号表を用いて可変長符号が割り当てられる。MPEG-2においては、ピクチャ単位で指定可能なパラメータ intra vlc format が 0 の場合、または非イントラマクロブロックの場合は、MPEG-1 と同一の符号表が用いられるが、パラメータ intra vlc format が 1 で、かつイントラマクロブロックの場合は、新たに定められた別の符号表を用いて可変長符号が割り当てられる。表 8 に、それぞれの可変長符号の符号長を run および level の値が 15 以下の範囲について示す。後者は、高レートで I ピクチャ(イントラマクロブロック)を符号化する場合を想定し、分散の大きな信号に適合した符号長分布となっている。

(b) Escape 符号

比較的出現頻度の低い run, level のペアに対しては、Escape 符号(6ビット)および run(6ビット)に続き、MPEG-1においては、level の絶対値が 127 までの場合は level を 8 ビット長で固定長符号化し、128 から 255 までの場合は 16 ビット長で固定長符号化する。MPEG-2においては、level の絶対値によらず、常に 12 ビット長で固定長符号化する。

MPEG-1においては、level の絶対値が 255 を超える値に対する符号が存在しないので、level はすべて 255(-255)にクランプされるのに対し、MPEG-2においては、発生し得るすべての level(絶対値が 2047 まで)に対して符号が存在する。

(c) End of Block (EOB)

ブロック内のスキャン順で、それ以降の係数値がす

表 8(a) run, level に対する可変長符号の符号長
(intra vlc format=0) (run ≤ 15, level ≤ 15)

level	run	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
15	14	17															
14	14	16															
13	14	16															
12	14	16															
11	13	16															
10	13	16															
9	13	16															
8	13	16															
7	11	14															
6	9	14															
5	9	13	14														
4	8	11	13	14													
3	6	9	11	13	13	14	17										
2	5	7	8	9	11	11	13	13	13	14	14	17	17	17	17	17	17
1	3	4	5	6	6	7	7	7	8	8	9	9	9	9	11	11	11
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	15	15

表 8(b) run, level に対する可変長符号の符号長
(intra vlc format=1) (run ≤ 15, level ≤ 15)

level	run	15	9	17													
15	9	17															
14	9	16															
13	9	16															
12	9	16															
11	9	16															
10	9	16															
9	8	16															
8	8	16															
7	7	14															
6	7	14															
5	6	9	14														
4	6	9	11	14													
3	5	8	9	13	13	14	17										
2	4	6	8	9	9	10	13	13	13	14	14	17	17	17	17	17	17
1	3	4	6	6	7	7	8	8	8	8	8	9	9	9	10	10	10
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	15	15

べて零となった場合、EOB(End of Block)と呼ばれる符号を出力し、そのブロックに対する可変長符号化を終了する。

(尾高)

3-2-5 レート制御とバッファ制御

1. ビット配分、レート制御と適応量子化

ビット配分、レート制御と適応量子化は、発生符号量を目標量に制御しつつ、同時に画質を最適化する重要な項目である。

特に MPEG ビデオ符号化方式においては、性質の異なる3種類のピクチャ(I, P, Bピクチャ)が存在するため、単一の送信バッファ容量を基にして量子化スケールをフィードバック制御する従来の方法では、MPEG ビデオ符号化方式の有する性能を最大限生かすことができない。MPEG ビデオ符号化方式に対するレート制御においては、特に以下の点を考慮する必要がある。

- (1) ピクチャタイプ毎に発生符号量が大幅に異なる。また、量子化スケールの変化に対する発生符号量の変化の割合も、ピクチャタイプ毎に大幅に異なる。
- (2) I, Pピクチャの量子化ノイズは、これを参照するP, Bピクチャへ伝搬するが、Bピクチャを参照することはないため、Bピクチャの量子化ノイズは伝搬しない。したがって、すべてのピクチャに対する量子化スケールを均一とすることが必ずしも全体の画質を最適化することにならず、この点を考慮した最適化が必要となる。

本節では、これらの点を考慮して MPEG-2 の TM5¹⁰⁾ で採用された方式を説明する。本方式は、方式提案時に高画質を達成した機関の方式を基本としたものであり、MPEG-1 の SM と比較して高画質を達成するものである。

本方式は、各ピクチャへのビット配分を行うステップ1、仮想バッファを用いたレート制御を行うステップ2、および視覚特性を考慮した適応量子化を行うステップ3の3つの階層から構成されている。

1.1 ステップ1

ステップ1では、GOP 内の各ピクチャに対する割当ビット量を、割当対象ピクチャを含め GOP 内でまだ符号化されていないピクチャに対して割り当てられるビット量 R を基にして配分する。この配分を GOP 内の符号化ピクチャ順に繰り返す。

その際、以下の2つの仮定を用いて各ピクチャへの

符号量割当を行っているのが特徴である。

(a) 各ピクチャを符号化する際に用いる平均量子化スケールコードと発生符号量との積は、画面が変化しない限り、ピクチャタイプ毎に一定値となると仮定する。

そこで、各ピクチャを符号化した後、各ピクチャタイプ毎に、画面の複雑さを示すパラメータ X_i, X_p, X_b (global complexity measure) を(18)式により更新する。このパラメータにより次のピクチャを符号化する際の量子化スケールコードと発生符号量の関係を推定できる。

$$X_i = S_i Q_i, \quad X_p = S_p Q_p, \quad X_b = S_b Q_b \quad (18)$$

ここで S_i, S_p, S_b はピクチャ符号化時の発生符号ビット量、 Q_i, Q_p, Q_b はピクチャ符号化時の平均量子化スケールコードである。また、初期値は目標レートである bit rate [bits/sec] を用いて(19)~(21)式で示される値とする。

$$X_i = 160 \times \text{bit rate} / 115 \quad (19)$$

$$X_p = 60 \times \text{bit rate} / 115 \quad (20)$$

$$X_b = 42 \times \text{bit rate} / 115 \quad (21)$$

(b) Iピクチャの量子化スケールコードを基準としたP, Bピクチャの量子化スケールコードの比率 K_p, K_b が(22)式に定めた値となる場合に、常に全体の画質が最適化されると仮定する。

$$K_p = 1.0, \quad K_b = 1.4 \quad (22)$$

上記の2つの仮定より、GOP 中の各ピクチャに対する割当ビット量 (T_i, T_p, T_b) は(23)~(25)式で示す値とする。

$$T_i = \max \left\{ \frac{R}{1 + \frac{N_p X_p}{X_i K_p} + \frac{N_b X_b}{X_i K_b}}, \right. \\ \left. \text{bit rate} / (8 \times \text{picture rate}) \right\} \quad (23)$$

$$T_p = \max \left\{ \frac{R}{N_p + \frac{N_b K_b X_b}{K_p X_p}}, \right. \\ \left. \text{bit rate} / (8 \times \text{picture rate}) \right\} \quad (24)$$

$$T_b = \max \left\{ \frac{R}{N_b + \frac{N_p K_p X_p}{K_b X_b}}, \right. \\ \left. \text{bit rate} / (8 \times \text{picture rate}) \right\}$$

$$\left. \text{bit rate}/(8 \times \text{picture rate}) \right\} \quad (25)$$

ここで N_p, N_b は GOP 内でまだ符号化されていない P, B ピクチャの枚数である。

すなわち、まず GOP 内の未符号化ピクチャのうち、割当対象となるピクチャとピクチャタイプの異なるピクチャについては、画質最適化条件のもとで、そのピクチャの発生する符号量が、割当対象ピクチャの発生符号量の何倍となるかを推定する。

次に、未符号化ピクチャ全体の発生する推定符号量が、割当対象ピクチャの何枚分の符号量に相当するかを求める。

割当対象ピクチャに対するビット量は、未符号化ピクチャに対する割り当てビット量 R を、この枚数で割ることによって与えられる。ただし、ヘッダ等に固定的に必要な符号量を考慮して、その値に下限を設定している。

このようにして求めた割当符号量を基にして、各ピクチャを符号化する毎に、GOP 内の未符号化ピクチャに対して割り当てられるビット量 R を(26)式で更新する。

$$R = R - S_{i,p,b} \quad (26)$$

また GOP の最初のピクチャを符号化する際には、(27)式により R を更新する。

$$R = \text{bit rate} \times N / \text{picture rate} + R \quad (27)$$

N は GOP 内のピクチャ数である。またシーケンスの最初での R の初期値は 0 とする。

1.2 ステップ 2

ステップ 2 では、ステップ 1 で求められた各ピクチャに対する割り当てビット量 (T_i, T_p, T_b) と実際の発生符号量を一致させるため、各ピクチャタイプ毎に独立に設定した 3 種類の仮想バッファの容量をもとに、量子化スケールコードをマクロブロック単位のフィードバック制御で求める。

まず、 j 番目のマクロブロック符号化に先立ち、仮想バッファの占有量を(28)～(30)式によって求める。

$$d_j^i = d_0^i + B_{j-1} - \frac{T_i(j-1)}{\text{MB cnt}} \quad (28)$$

$$d_j^p = d_0^p + B_{j-1} - \frac{T_p(j-1)}{\text{MB cnt}} \quad (29)$$

$$d_j^b = d_0^b + B_{j-1} - \frac{T_b(j-1)}{\text{MB cnt}} \quad (30)$$

d_0^i, d_0^p, d_0^b は各仮想バッファの初期占有量、 B_j はピクチャの先頭から j 番目のマクロブロックまでの発生ビット量、MB cnt は 1 ピクチャ内のマクロブロック数である。

各ピクチャ符号化終了時の仮想バッファ占有量 ($d_{\text{MB cnt}}^i, d_{\text{MB cnt}}^p, d_{\text{MB cnt}}^b$) は、それぞれ同一のピクチャタイプで、次のピクチャに対する仮想バッファ占有量の初期値 (d_0^i, d_0^p, d_0^b) として用いられる。

次に、 j 番目のマクロブロックに対する量子化スケールコードを(31)式により計算する。

$$Q_j = \frac{d_j \times 31}{r} \quad (31)$$

r はリアクションパラメータと呼ばれるフィードバックループの応答速度を制御するパラメータであり、(32)式で与えられる。

$$r = 2 \times \text{bit rate} / \text{picture rate} \quad (32)$$

なお、シーケンスの最初における仮想バッファ初期値は(33)式で与えられる。

$$d_0^i = 10 \times r / 31, \quad d_0^p = K_p d_0^i, \quad d_0^b = K_b d_0^i \quad (33)$$

1.3 ステップ 3

ステップ 3 では、ステップ 2 で求められた量子化スケールコードを、視覚的に劣化の目立ちやすい平坦部でより細かく量子化し、劣化の比較的目立ちにくい絵柄の複雑な部分でより粗く量子化するように、各マクロブロックごとのアクティビティによって変化させている。

アクティビティは、予測誤差でなく原画の輝度信号画素値を用い、フレーム DCT 符号化モードにおける 4 個のブロックとフィールド DCT 符号化モードにおける 4 個のブロックとの合計 8 個のブロックの画素値を用いて(34)～(36)式で与えられる。

$$\text{act}_j = 1 + \min_{sblk=1,8} (\text{var } sblk) \quad (34)$$

$$\text{var } sblk = \frac{1}{64} \sum_{k=1}^{64} (P_k - \bar{P})^2 \quad (35)$$

$$\bar{P} = \frac{1}{64} \sum_{k=1}^{64} P_k \quad (36)$$

ここで P_k は原画の輝度信号ブロック内画素値である。(34)式において最小値を採るのは、マクロブロック内の一部だけでも平坦部分のある場合には量子化を細かくするためである。

さらに、(37)式により、その値が 0.5～2 の範囲をとる正規化アクティビティ $N\text{act}_j$ を求める。

$$N\text{act}_j = \frac{2 \times \text{act}_j + \text{avg act}}{\text{act}_j + 2 \times \text{avg act}} \quad (37)$$

avg act は直前に符号化したピクチャでの act_j の平均値である。

視覚特性を考慮した量子化スケールコード $m\text{quant}_j$ はステップ 2 で得られた量子化スケールコード Q_j を基に(38)式で与えられる。

$$mquant_j = Q_j \times Nact_j \quad (38)$$

1.4 制限事項

TM5で採用された本方式には、以下の制限のあることが知られている。したがって、本方式を基本として実際の制御を行う場合は、これらの制限に対する対策が必要となる。

- ・ステップ1はシーンチェンジに対応できない。またシーンチェンジ後にステップ3でavg actが間違った値となる。
- ・次節で説明するVBVに対する拘束条件を満足する保証がない。

2. VBV

VBV (Video Buffering Verifier) は、エンコード出力に接続される仮想デコードモデルであり、このモデルに含まれる受信バッファ (VBV バッファ) の占有量に対する拘束条件によって、ビットストリームに対する制限を規定するものである。

VBVの目的は、あくまでこのモデルによってビットストリームに対する規定を加えることであり、現実のデコードの実現方式を規定するものではない。

VBVは、以下の理想条件のもとで動作することを仮定している。

- ・エンコードとVBVは完全に同期して動作する。
- ・各ピクチャの復号は瞬時に行われ、各ピクチャに対するデータはVBVバッファより瞬時に抜き取られる。

2.1 VBVに対するデータ入力

当初、VBVは固定レート (CBR) データのみに対する規定であったが、システム規格と併せた検討の結果、可変レート (VBR) データに対してもVBVによる規定を行うことになった。また、様々な応用に対応するため、VBVに対するデータの入力は以下に示す2種類の方法で規定できるようにした。いずれの場合も、そのピークレート $R_{max}(b/s)$ が、シーケンスヘッダに含まれるパラメータビットレートで指定される。

2種類の方法の区別は、ピクチャヘッダに含まれるパラメータ vbv delay の値によってなされる。

(1) vbv delay が 0xFFFF でない場合

各ピクチャの vbv delay が、すべて最大値 0xFFFF でない場合、vbv delay は各ピクチャの picture start code が VBV バッファに入力されてから復号されるまでの時間を示している。この場合、VBV に対するデータの入力レートは、ピクチャデータ単位で可変レートとなる。 n 番目のピクチャデータが VBV バッファに入力されるレート $R(n)(b/s)$ は (39)

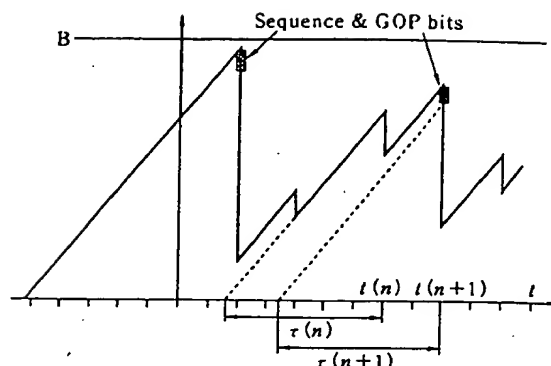


図 16 VBV バッファ占有量の変化 (CBR, vbv delay $\neq 0xFFFF$)

式で定められる。

$$R(n) = d_n^* / (\tau(n) - \tau(n+1) + t(n+1) - t(n)) \quad (39)$$

ここで、 d_n^* は n 番目のピクチャに対する picture start code の直後から $(n+1)$ 番目のピクチャに対する picture start code までの符号量、 $\tau(n)$ は n 番目のピクチャに対して設定されたパラメータ vbv delay の値、 $t(n)$ は n 番目のピクチャに対するデータが VBV バッファより抜き取られる時刻である。固定レート (CBR) データに対しては、 $R(n)$ は n によらず一定の値となる。

図 16 に VBV バッファ占有量の変化を示す。初期状態において占有量は零であり、最初の picture start code の入力時点からパラメータ vbv delay の時間後に復号を開始する。引き続き定められた復号時間間隔で各ピクチャの復号と VBV バッファの検証を継続する。

(2) vbv delay が 0xFFFF の場合

各ピクチャの vbv delay がすべて最大値 0xFFFF の場合、VBV に対するデータの入力レートは、VBV バッファの占有量で決定され、VBV バッファに空きがある時には R_{max} 、空きのない場合は零となる。

初期状態において VBV バッファ占有量は零であり、バッファに空きがなくなった時点で復号を開始する。引き続き、定められた復号時間間隔で各ピクチャの復号と VBV バッファの検証を継続する。

2.2 VBV バッファに対する検証

(1) 各ピクチャの復号時において、ピクチャデータを VBV バッファから瞬時に抜き取る直前に VBV バッファがオーバフローしてはならない。

(2) ピクチャデータを VBV バッファから瞬時に抜き取った直後における VBV バッファのアン

ダフローに関しては、パラメータ low delay で指定される低遅延モードの場合と、そうでない場合とで検証内容が異なる。

ここで低遅延モードとは、B ピクチャの使用を禁止し、符号化・復号化時におけるピクチャの並べ替えによって生じる遅延を防ぐとともに、送受信バッファでの遅延を少なくするため、符号化ピクチャのスキップ(駒落し)を許しているモードである(3-2-8を参照)。

低遅延モードでない場合には、VBV バッファからピクチャデータを瞬時に抜き取った直後に VBV バッファがアンダフローしてはならない。

低遅延モードの場合には、仮りにピクチャデータを抜き取ると、VBV バッファがアンダフローする状況となることを許す。この場合には、その時刻でピクチャデータを抜き取らず、2 フィールド間隔毎にすべてのピクチャデータが到達したかどうかのチェックを繰り返し、初めて到達した時点でピクチャデータを抜き取る。

したがって、低遅延モードの場合も含め 1 ピクチャのデータ量は VBV バッファサイズより小さくしなければならない。
(尾高)

3-2-6 スケーラビリティ

・スケーラビリティの概念

MPEG-2 には、スケーラビリティと呼ばれる機能が導入された。このスケーラビリティは、ビットストリームを部分的に復号することによって SNR (Signal to Noise Ratio)、空間解像度、時間方向の解像度などを段階的に可変とする機能である。スケーラビリティの実現には、階層符号化と呼ばれる手法を用いている。

・スケーラビリティのメリット

スケーラビリティは、いずれも階層符号化を行うため、非常に複雑な処理を要する。しかし、伝送帯域の限られた状況では、単純にそれぞれの階層を別々に符号化して伝送する手法(サイマルキャストと呼ばれる)に比較して、少しでも符号化効率を向上させ、符号化に必要なヘッダ情報やサイド情報の重複を避けることが重要となる。スケーラビリティは、各階層間でのこれらの情報の相互利用を行うことによって符号化効率を向上している。

いずれの方式も、現在すぐに実装・実用化することは困難に見えるが、将来の伝送・記録方式として注目されている。

・スケーラビリティとプロファイル

MPEG-2 で定義されているメインプロファイル以上の SNR/空間/High プロファイルは、スケーラビリティを実現するために階層符号化をサポートする。それぞれで採用した技術と機能性は以下の通り。

SNR スケーラブルプロファイル

符号化手法：SNR スケーラビリティ

実現される機能：グレースフルデグラデーション
空間スケーラブルプロファイル

符号化手法：空間スケーラビリティ

実現される機能：異なる解像度の信号間での互換性

High プロファイル

符号化手法：SNR スケーラビリティ、空間スケーラビリティ

実現される機能：SNR スケーラビリティと空間スケーラビリティの結合

上記した各プロファイルにおける符号化手法とその機能の詳細に関して、以下に説明する。

1. SNR スケーラブルプロファイル

1.1 概念と応用分野

SNR スケーラブルプロファイルは、後述する SNR スケーラビリティと呼ばれる技術を採用し、グレースフルデグラデーションという機能を提供する。グレースフルデグラデーションは、例えば電波の受信強度に応じた緩やかな画質の劣化を実現する。一般に、通常のデジタル放送の問題点として、電波の受信強度がある規定値以下になった場合、これを境界として全く放送が受信できなくなることが挙げられる。しかし、放送に SNR スケーラビリティを適用した場合は、画像の品質が緩やかに劣化し、現実的な放送範囲の拡大を達成し、なおかつ隣接した放送領域との境界付近に生じる干渉を抑えることができる。

1.2 SNR スケーラビリティの符号化方法

SNR スケーラビリティは、周波数領域(ここでは DCT 係数)の階層符号化を実現する。基本的な考え方としては、ベースとなる画質(基準となる低 SNR)を提供する信号を高い優先度の伝送路で送り出し、空間